Agriculture and Food

AIMS Agriculture and Food, 9(4): 980–1003.
DOI: 10.3934/agrfood.2024053
Received: 04 June 2024
Revised:18 September 2024
Accepted: 26 September 2024
Published: 12 October 2024

http://www.aimspress.com/journal/agriculture

Research article

Crop yield prediction through machine learning: A path towards sustainable agriculture and climate resilience in Saudi Arabia

Mohammad M. Islam[1], Majed Alharthi[1,*], Rotana S. Alkadi[1], Rafiqul Islam[2], Abdul Kadar Muhammad Masum[3]

[1] Department of Finance, College of Business, King Abdulaziz University, Rabigh, Saudi Arabia
[2] School of computing mathematics and engineering, Charles Sturt University, Australia
[3] Department of Software Engineering, Daffodil International University, Bangladesh

* **Correspondence:** Email: mdalharthi@kau.edu.sa; Tel: +966569943471.

Abstract: This study aimed to explain the crop yield prediction system as a way to address the challenges posed by global warming and climate change in Saudi Arabia, while also taking into account socio-economic factors. Machine learning models were trained using crop yield prediction data to provide recommendations for future crop production. Climate change poses significant challenges, with rising temperatures and extreme weather events being increasingly evident. Agriculture, contributing 14% of greenhouse gas emissions, plays a crucial role in exacerbating this issue. This study introduced a crop yield prediction system leveraging machine learning models trained on comprehensive datasets. Recommendations derived from these models offer insights into optimal crop rotation strategies, particularly relevant for regions like the Kingdom of Saudi Arabia. Collaboration between farmers and governments, informed by data-driven approaches, is crucial in this endeavor. Utilizing a customized dataset, this study analyzed a machine learning model performance and identified optimal hyperparameters. XGBoost ensemble emerged as the top performer with an R^2 score of 0.9745, showcasing its potential to advance crop yield prediction capabilities. By integrating machine learning into agricultural decision-making processes, stakeholders aim to enhance crop production and soil health and contribute to climate change mitigation efforts. This collaborative effort represents a significant step toward sustainable agriculture and climate resilience in Saudi Arabia.

Keywords: climate change; machine learning; precision agriculture; Saudi Arabia; sustainability; sustainable finance

1. Introduction

Climate change poses one of the most significant challenges confronting humanity. The repercussions of climate change are becoming more obvious day after day [1]. The average worldwide temperature and precipitation levels are predicted to rise due to climate change (by 1.4 to 5.8 °C by 2100) [2]. Droughts, storms, flooding, and wildfires are becoming increasingly severe and frequent [3]. If nations with limited water resources cannot adapt, there will be significant impacts on both the quantity and quality of water, posing serious risks to public health. Global ecosystems are evolving, as are the natural and agricultural resources on which humans rely. According to the 2018 UN report on climate change, catastrophic consequences await if worldwide greenhouse gas emissions are not halted within the next thirty years [4]. Yet, each year, these emissions increase.

Approximately 14% of greenhouse gas (GHG) emissions are attributed to agriculture [5]. As plants absorb carbon dioxide from the atmosphere, this statistic may appear unexpected. However, modern industrial agriculture encompasses more than merely plant cultivation. Initially, to release stored carbon, the area undergoes deforestation. Subsequently, tilling disrupts soil-dwelling organisms that are crucial for sequestration by exposing topsoil to the atmosphere and releasing carbon trapped in soil aggregates. Lastly, nitrogen-based fertilizers must be reintegrated into the system due to the depletion of soil nutrients resulting from these farming practices. Approximately 2% of the world's energy is used in the vast quantities required to synthesize these fertilizers [6]. Furthermore, some of this nitrogen is transformed into nitrous oxide [7], a greenhouse gas that is around 300 times more powerful than CO_2, while the remainder is absorbed by plants or held in the soil.

Farmers are increasingly seeking sophisticated technologies that enable them to labor at scale while adapting to their land's demands. This strategy is commonly referred to *precision agriculture*. Precision agriculture can minimize soil carbon emissions while increasing crop yields. There are numerous other ways in which machine learning might contribute to precision agriculture. Within the field of artificial intelligence (AI), machine learning is the study and creation of computer systems that learn automatically from data and experience. Machine learning allows computers to learn from data and develop predictions or judgments without the need for explicit programming. Irrigation systems can conserve water and reduce pests that thrive in high-moisture environments [8]. Machine learning may also aid in disease diagnosis, weed identification, and soil sensing [9–11]. Machine learning may aid in agricultural production prediction [12] and macroeconomic models, assisting farmers in determining what to plant at the start of the season [13].

Climate change mitigation includes lowering emissions as well as adaptation (preparing for unavoidable repercussions). Both are intricate concerns. To reduce GHG emissions, improvements must be made to electrical systems, transportation, structures, industry, farming, and land use practices. Adaptation necessitates endurance and disaster management planning, as well as knowledge of climate and severe occurrences. Such a wide range of issues might be viewed as an opportunity—there are many methods to make a difference. In the past few years, AI has emerged as a significant tool for technological advancement. Despite the growing number of initiatives using machine learning and AI to address social and global problems [3], there is still a need for a concentrated effort to determine how these technologies may be best deployed to combat climate change. Many machine learning practitioners want to intervene but are unsure how. On the other hand, several sectors have started actively soliciting feedback from the AI community.

Climatic factors include rainfall and temperature. These abiotic components, together with pesticides and soil, are environmental factors that influence plant growth and development. Rainfall has a dramatic effect on agriculture. For this project, information on total rainfall per year was gathered from the World Data Bank. Crop rotation, which involves growing different crops in the same area at different times, increases soil health as many plants utilize different mixes of nutrients in the soil. Rotating crops helps to maintain an even amount of nutrients, which is key for optimal soil health. Studies have shown that these practices not only capture more carbon dioxide from the air but also help farmers improve their crops in terms of health and yield, enabling them to grow more nutrient-dense food.

This research answers the question of how can machine learning-based crop yield prediction systems aid in optimizing crop rotation and mitigating climate change impacts in Saudi Arabia. Reduced tillage intensity or no tillage at all and plant residue preservation have the potential to help mitigate climate change by lowering soil GHG emissions [14]. Hence, this study proposes a crop yield prediction system aimed at addressing the challenges posed by global warming and climate change, while also taking into account socio-economic factors. Machine learning models were trained using crop yield prediction data to provide recommendations for future crop production. These recommendations offer valuable insights for implementing crop rotation strategies tailored to specific soil conditions. Emphasizing the Kingdom of Saudi Arabia, collaboration between farmers and the government can facilitate the adoption of reformed agricultural practices and policies informed by the proposed system. The main contribution of this study lies in the development and training of a resilient machine-learning model. Its purpose is to accurately forecast crop yields for forthcoming seasons, facilitating farmers in seamlessly rotating their crops between seasons without the need for tilling. This collaborative effort aims to enhance crop production and soil health and contribute toward mitigating global warming.

The remainder of the paper is structured as follows. Relevant works are discussed in Section 2. The dataset used is described in Section 3. The methodology of the study is presented in Section 4, along with an explanation of the regression and data preparation methods applied. The acquired regression findings are shown, and their reliability is justified in Section 5. The study's theoretical and practical implications are presented in Section 6. In conclusion, Section 7 offers insights into the study, a summary of the whole work, and suggests areas for further investigation.

2. Literature review

Since the beginning of the 1980s, because of the significant rise in carbon dioxide levels and other trace gases in the atmosphere, several climatologists projected that future decades would see major global warming. As an initiative to deal with such a problem, the United Nations Environmental Program (UNEP) and the World Meteorological Organization (WMO) jointly established the Intergovernmental Panel on Climate Change (IPCC) to examine and evaluate the scientific, technical, and socioeconomic information pertinent to the understanding of human-created climate change, its possible effects, and options for adaptation and mitigation [15].

2.1. Impact of climate change on agriculture

The impact of climate change and variability on agriculture has been well documented in the literature. The prevalent belief is that variations in temperature and precipitation will cause changes in the land and water regimes, which will then have an impact on agricultural output [16].

Trends of increasing average temperature and more volatile rainfall patterns have been found by the National Academy of Science (2001). Additional evidence from [17] indicates that global climate systems are changing more quickly than anticipated, increasing the likelihood of more severe and abrupt changes. For the main crops (wheat, rice, and maize) in temperate and tropical regions, climate change without adaptation is likely to have an adverse effect on production, as local temperature has increased by 2 °C or more than late 20th century levels, according to the Fifth Assessment Report of IPCC (2014) [18]. Another prediction suggests that the average global temperature will rise by 2.4 °C between 1990 and 2100, with a 95% possibility that the increase will be between 1.0 °C and 4.9 °C [19]. Other studies expect that the average global temperature may rise by 0.3 to 1.3 K during the coming 30 years [20], which will ultimately have an adverse impact on the agricultural output.

It is believed that most high temperatures and global warming that we are going to experience during the next 40 years will be due to emissions that have already happened. This means that, in the longer term, the pace and degree of global warming heavily rely on currently occurring and near-future emissions. There is a greater than 50% chance that the increase in temperature over time will exceed 60 °C. The impacts of climate change amount to 5% of the global GDP, which grows regionally by up to 20%, and such climate change is anticipated to represent an annual loss in the future [21].

More importantly, several studies indicate that the impact of global warming and climate change is different from one context to another, making it important to focus on a particular context. In this study, the focus is on Saudi Arabia.

2.2. The impact of climate change on agriculture: the case of Saudi Arabia

Saudi Arabia has been emitting more CO_2 on a consistent basis, which contradicts the global mandate for the mitigation of GHG for sustainable development. The United Nations Climate Change Conference, or COP 21, held in 2015, strengthened the global movement to keep global warming to less than 2 °C compared to pre-industrial levels. This challenging objective necessitates a significant reduction in GHG emissions, especially CO_2. According to the most recent estimates from 2022, the Kingdom is currently one of the largest emitters of greenhouse gases in the world, and its emissions reached more than twice the G20 average [22]; given the rate at which emissions are increasing, this is likely to rise quickly.

The climate of Saudi Arabia is hot, arid, and harsh, and any slight change in climate can have a significant impact on agriculture and water resources and availability. Keeping in mind the importance of this possible danger, an assessment study was conducted in [23]. Based on the analysis of four decades' data, authors reported that a decrease in precipitation and an increase in temperature could have a major adverse impact on water supplies and agriculture.

Saudi Arabia is known and considered to be one of the top three date-producing nations in the world. According to the latest report (2020) of The Ministry of Environment, Water, and Agriculture (MEWA), date manufacturing in Saudi Arabia reached 1.54 million tons in 2020. Nevertheless, despite the enormous government support and awareness of date palm cultivation in the country, exports of dates have not reached the expected level because the date-productivity level is still low compared with other date-producing countries [24]. This decline could be due to several factors, such as insect pests and plant diseases as well as environmental stress factors including salinity, drought, and temperature extremes due to climate change. Hence, it can be anticipated that climate change will have a significant influence on Saudi Arabia's agriculture and food production, particularly through reduced water availability and direct effects on crop yields. For instance, during the 2010 season, several farmers recognized the strange early flourishing of date palms as a direct result of climate change [25]. However, to deal with such changes, according to the regional competitive advantage, the Kingdom considered the optimization of the cropping model to adapt to the unfavorable effects of climate change [26].

With the current harsh climate challenges and water scarcity in the production of many important crops in Saudi Arabia, such as dates, and to penetrate the business of sustainable agriculture, the Kingdom's agricultural sector is going through a period of great change, trying to develop new technologies and farming techniques that are suitable for dry climates, as reported by the Sustainable Agriculture Development Research Centre (SADRC) [27]. It is important to note that, since dates are one of Saudi Arabia's most important crops, it is urgent to investigate further research options in this area [28]. More importantly, given the high sensitivity of the agricultural sector to climate change, applying adaptation strategies rather than just mitigation ones is crucial to the country [29].

2.3. Climate change adaptation strategies

Many alternative working definitions of climate adaptation have naturally emerged as a result of the broad and diversified interest in adaptation. Adaptation is defined by IPCC as "initiatives and measures to reduce the vulnerability of natural and human systems against actual or expected climate change effects" [17] and "the adjustment in natural or human systems in response to actual or expected climatic stimuli or their effects, which moderates harm or exploits beneficial opportunities" [30].

2.3.1. Adaptation techniques to climate change in Saudi Arabia

The adaption techniques that are suitable to climate change worldwide have been investigated by some researchers, while other studies have recommended some techniques that can be used in Saudi Arabia in particular.

The hydroponic green fodder production technique has been reported to only need between 2% and 10% of the water needed to grow the same amount of food in a soil culture [31]. Furthermore, compared to fodder produced in field circumstances, the hydroponic technique requires only 3%–5% of water to produce the same amount [32]. Hence, a hydroponic system is a promising technique to produce food using a lower amount of water, with particularly great potential in countries where there is water scarcity in agricultural production, such as Saudi Arabia.

Possible adaptation strategies to achieve food security and self-sufficiency have been examined by [33]. Authors suggested that the agricultural industry and the government should support and be willing to implement innovative water-saving technologies such as hydroponics, greenhouse farming, and seawater harvesting. The authors also emphasized the significance of the extension agent's role in promoting these technologies and educating farmers about effective water-saving agricultural techniques. In [23], the potential impacts of climate change and global warming on Saudi Arabia's agricultural industry and water supply were investigated using general circulation models (GCM). The findings suggested that a rise in temperature and a fall in precipitation might be damaging for the Kingdom's agricultural and water supply. Some adaptation strategies to climate change that could be used in Saudi Arabia were also suggested. First, the use of rechargeable or replenished deep aquifers, which could be done through well-built holes that link reservoirs with the water tables of the aquifers. Second, more effective irrigation methods and systems including sprinklers and drip irrigation. To support such an adaptation method, they referred to a practice from some areas in the US with similar climate to Saudi Arabia (Arizona and Nebraska) where the usage of groundwater pumping for irrigation has been replaced with more efficient irrigation systems [34]. Third, an increase in the usage of greenhouse farming, as the total percentage of such greenhouses to total agricultural production is considered to be very low in the Kingdom. Greenhouses are very effective in using much less water due to lower evaporation rates. Fourth, stopping overgrazing was suggested, as it leads to soil erosion and a reduction in soil quality. Authors also recommended applying stricter water usage policies in the Kingdom. The final recommendation was that more research on climate change and agriculture is needed in Saudi Arabia, which increases the importance of the current study. The findings of our study are expected to support the country's sustainability and provide policymakers with some recommendations to reduce the impact of climate change on their agricultural productions; all of which will support the achievement of Vision 2030 goals, which have agricultural sustainability at its top list.

[29] evaluated how climate change affects agricultural crop production using historical data in Saudi Arabia. The main findings show that an increase in the temperature by 1 °C results in a 7%–25% reduction in crop yields. Due to this high risk of future yield production in the Kingdom, they identified some adaptation strategies. Particularly, as an agronomical and technological adaptation, they suggested planting and growing drought-tolerant crops (such as planting dates palms), using mixed cropping and intercropping, as well as the use of sustainable agriculture. They also suggested the use of micro-irrigation techniques including drip and sprinkler irrigation, which will help in conserving the water and using it more efficiently. They also recommended that the Kingdom urgently needs to adopt modern agricultural technology such as hybrid varieties, less water-intensive crop varieties, and breeding for stress tolerance. Additionally, the authors noted the high need to provide farmers with better loans/credit delivery and crop insurance to encourage them to implement climate adaptation strategies.

2.3.2. Socioeconomic factors of farmers and their impact on climate change in Saudi Arabia

There is a serious need to increase the awareness of sustainable agriculture practices among farmers [35]. Therefore, [36] suggested that the introduction of successful sustainable agriculture programs would need the support of the Extension Service and its employees. [37] also recognized the need to develop sustainable agriculture programs to educate and train farmers on sustainable agriculture technologies. Sustainable agriculture is a production system that has the possibility of addressing many constraints and issues faced by resource-poor farmers while also being widely accepted by society. It refers to how agriculture can, over time, contribute to the general welfare of the communities by ensuring that they have access to enough food and other goods and services in ways that are financially feasible, socially acceptable, and environmentally sound. Social sustainability, on the other hand, refers to the quality of life of both farmers and the communities in which they operate [38].

From a socioeconomic perspective, the focus primarily lies on the examination of the social, political, and economic aspects pertinent to specific individuals or social groups within the society [39]. Generally, the socioeconomic approach focuses on identifying the individuals' and communities' adaptive capabilities based on their inner characteristics such as education, gender, health status, wealth, access to technology and information, political power, and formal and informal (social) capital. The differences in strength levels between localities are caused by variations in these elements concerning climate change. For instance, agriculture is a primary source of income for farmers. When agricultural productivity and output are reduced, their earnings also decrease. Hence, the sustainability of the farmers' social and economic systems is impacted by all of these elements, either directly or indirectly.

As documented by the 2030 strategic plan of the Ministry of Environment, Water, and Agriculture (MEWA), there is a significant lack of knowledge in Saudi Arabia regarding how climate change affects people's livelihoods and the need to build local capacity to adapt to it [40]. Some studies aimed to fill this gap. For instance, [41] took a bottom-up approach by investigating how farmers in Southern Saudi Arabia (Jazan) with different characteristics perceive climate change. They also investigated the farmer's perception of the role of extension services in improving their capacity to adapt to climate change. Based on interviews with 164 farmers in the Jazan region, they found that farmers who believe that climate change is due to natural change are more than those who believe that climate change is due to human activities. More than 70% of farmers are concerned about insects and the prevalence of weeds on their farms. They also found that farmers who are more inclined to believe in climate change also agree with the importance of extension services in fostering capability. In terms of the variables that influence the farmers' beliefs on climate change, [42] found that loan availability, use of extension services, membership in agricultural cooperatives, age, soil fertility, and farm size all had a big impact on farmers' views on climate change. However, their results show that loan availability is the only statistically significant factor that explains the variation in farmers' concerns. They suggested that there is a high demand for actions to be taken to enhance farmers' capability to manage climate variability. [43] empirically investigated how different socioeconomic indicators influenced farmers' worries and highlighted several measures to enhance community capacity for effective adaptation in Jazan, Saudi Arabia. Results from ordered logit models revealed a significant association between some different variables. The farmers' level of income and age were significantly and positively associated with their concerns about drought. The level of income and education of farmers had a negative association with their worries about insect infestations. A significant inverse association between income level and farmers' worries about a rise in disease frequency was also evident. Surprisingly, farmers who have access to credit facilities reported higher concern regarding all three effects of climate change-related concerns (drought, insect infestation, and diseases). In contrast, the availability of information about climate change greatly lowered farmers' worries about increased disease occurrence and insect infestation. Aside from climate change worries, farmers identified three key capacity-building initiatives as effective for enhancing climate change adaptation, including the use of the information and communication technologies (ICT) tools to raise farmers' awareness of climate change issues and pertinent adaptation practices, capacity development of extension personnel to improve their knowledge, and connecting smallholder farmers with agricultural researchers to create farm-based climate adaptation strategies [43]. [44] investigated the farmers' socioeconomic characteristics impact on their perceptions and awareness about the negative effects of pesticides on the environment. Applying a questionnaire to 204 farmers in Dawadmi Province of Saudi Arabia, they found that 5% of farmers depend less on agricultural extension and instead look for information from other reliable sources. The farmers are eager to learn about the harmful impacts of pesticides on the ecosystem. The study additionally demonstrated that spraying with axis sprayers or portable sprayers is the most typical method of applying pesticides. The authors suggested that there is a high need for establishing extension programs on proper and secure pesticide handling and application techniques. Moreover, similar to [43], they recommended a connection between the farmers and the agricultural researchers.

[44] investigated the climate change and the global warming impact on date production in the Al-Hassa region of Saudi Arabia, which is one of the largest date-producing regions with around 3 million date palms. Based on numerous models [i.e., general circulation models (GCM)], regional climate models (RCM), and emission scenarios on the climate in Al-Hassa, the author found that climate change effects will put pressure on date production in the Al-Hassa region by the end of the 21st century (2071–2100). [45] also argued that water shortage and a labor force lacking knowledge and skills are the two key resources that prevent farms from surviving. The author suggested that to preserve date production in Al-Hassa, adaptive and mitigating action must be taken. Particularly, [45] suggested that new technology, more productive farming methods, financing for ongoing research and development, and more effective regulations must be implemented. It was also recommended that a multilevel governance framework should be in place to promote procedural equity, helping to overcome institutional and cultural barriers leading to a higher rate of adoption. All stakeholders, including farmers, the community, and all governmental levels, should be involved. Obstacles and limits related to the environment that prevent the adoption of new practices may exist, but they can be overcome with enough assistance, education, and awareness [45].

[41] examined the farmers' awareness of environmental agri-environmental legislation that intends to protect the environment. They particularly tested Saudi farmers' perceptions of several components of environmental protection legislation—knowledge of legislation, penalties for violations, and the negative impact of non-compliance. The study was based on interviews with 312 farmers of Al Kharj governorate of Riyadh region in the center of Saudi Arabia. They found that around 20% of the farmers have an awareness of the environmental regulations. Also, they found that among the various types of regulations, farmers' knowledge of pesticide regulations received the lowest score. Additionally, knowledge of penalties had a lower score than knowledge of legislation and negative environmental impacts. They recommended that policymakers should address the issue of low legislative awareness by focusing on each component of legislation. Particularly, they recommended that integrated extension messages should be developed, in which the meaning of each piece of legislation, the penalties associated with it, and the environmental risks associated with non-compliance are explained. [31] criticized the massive production of wheat in the Kingdom in recent years, highlighting that such a surplus in the production of wheat not only drained the Kingdom's water resources, which are primarily drawn from non-renewable aquifers, but also necessitated the use of massive amounts of chemical fertilizers to boost yields, which was not sustainable. They claimed that this type of agricultural production does not meet the definition of sustainability. Hence, they recommended that only segments that are sustainable and capable of improving crop yields and rural livelihood of farming communities should be prioritized by the Kingdom. Particularly, they recommended that greenhouse agriculture, aquaculture, poultry, dairy production, and technology to produce more diverse crops with less water must all be prioritized. More importantly, they recommended that there is a growing need to educate farmers in Saudi Arabia in the use of agricultural input as efficiently as possible and view natural resources such as land and water as non-renewable.

The literature review shows that researchers recognized the importance of applying adaptation strategies in Saudi Arabia to mitigate the impact of climate change on agriculture. However, most of the studies only provide recommendations for some adaptation techniques and very limited studies have empirically examined their effects. Moreover, studies showed that the socioeconomic factors of farmers and their awareness of climate change do have an impact. However, the socio-economic factors are very wide and not all factors have been investigated. More importantly, studies that examined such an aspect are very limited in the first place. Hence, the current study aims to shed light on these areas. Moreover, since the Vision of 2030 is giving significant attention to the growth of the non-oil sector, this study is considered to be an essential opportunity for farmers and stakeholders (i.e., government) to jointly work on the reformation of agricultural practices and policies.

3. Dataset description

A standardized benchmark dataset for agricultural research has yet to emerge, primarily due to regional variations, crop types, climates, and irrigation techniques [46,47]. Hence, researchers in this field employ datasets customized to their research objectives, aligning precisely with the specific purposes of their studies. Therefore, in this study, the crop yield prediction [48] dataset was used, where pesticides and yield data were collected by FAO (Food and Agriculture Organization), and rainfall and average temperature were sourced from the World Bank Open Data. The final dataset was cleaned and merged to include pesticides, yield, rainfall, and average temperature in one file. This dataset comprises 10 of the most consumed crops worldwide: maize, potatoes, rice/paddy, sorghum, soybeans, wheat, cassava, sweet potatoes, plantains, and yams [48]. Among these, following thorough cleaning and integration, the finalized dataset spans 23 years, encompassing 101 nations from 1990 to 2013.

Table 1 displays the variables used in this dataset along with brief descriptions. The variable "ha/hg yield" was utilized as the label. The data in this study was collected from 101 countries. It encompasses numerous Middle Eastern and North African (MENA) nations, such as Bahrain, Saudi Arabia, Iraq, Egypt, Turkey, Libya, Algeria, and Morocco. These countries have been chosen for their weather characteristics, close to Saudi Arabia's weather. Figure 1 illustrates that there is no noticeable correlation among the columns of the dataset variables. Table 2 shows a sample of the crop yield prediction dataset, where each row represents a specific entry in the dataset with data recorded for a particular combination of variables.

Table 1. Variables description of the dataset.

Variable	Description	Source
Area	Area based on different nations	[49,50]
Item	Planted crops name	[49,50]
Avg temp	Average temperature in Celsius	[49,50]
Pesticides tons	Pesticides used in tons	[49]
Average rainfall per year	Average rainfall per year	[50]
Year	Planted crops time (1990–2013)	[49,50]
hg/ha yield	Production value of crops yield in hectogram per hectare (Hg/Ha)	[49]

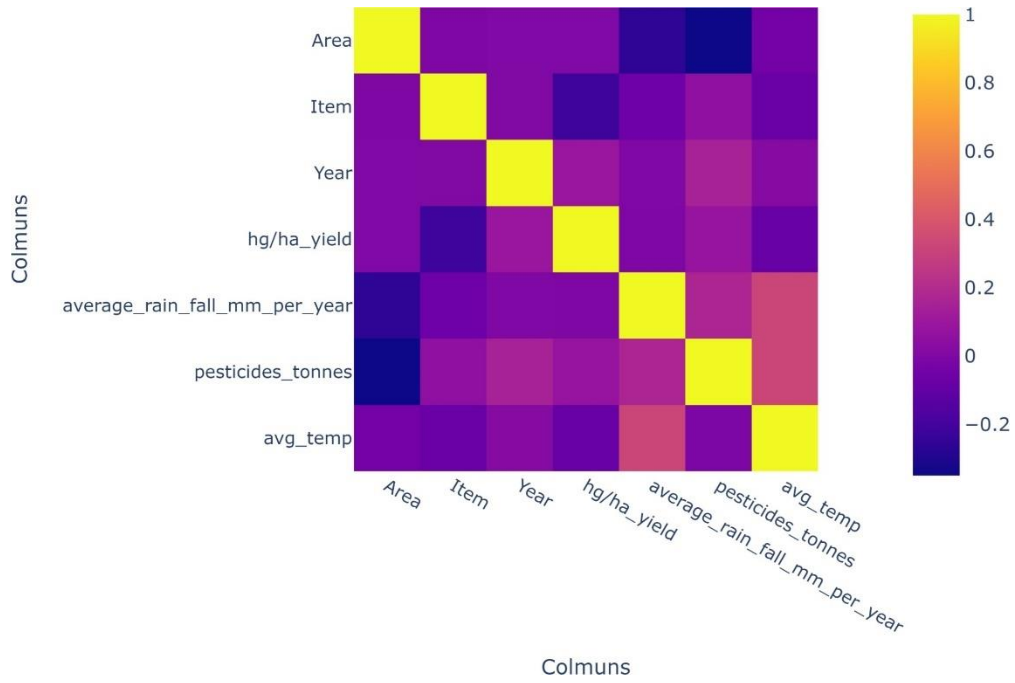


Figure 1. Correlation between dataset columns.

Table 2. A sample of crop yield prediction dataset.

Index	Area	Item	Year	Hg/ha_yield	Ave. rainfall. mm/year	Pesticides tons	Ave. Temp
24320	Saudi Arabia	Potatoes	2012	253,233	59.0	4996.2	26.83
24321	Saudi Arabia	Sorghum	2012	26,056	59.0	4996.2	26.83
24322	Saudi Arabia	Wheat	2012	59,254	59.0	4996.2	27.02
24323	Saudi Arabia	Maize	2013	61,024	59.0	5412.5	27.57
24324	Saudi Arabia	Potatoes	2013	256,547	59.0	5412.5	27.57

4. Methodology

Figure 2 illustrates the general approach for describing the impact of climate change on crop yield prediction. First, the dataset was extracted from FAO and World Data Bank. Furthermore, the Saudi Arabia area data were extracted from the primary dataset [48] and underwent a thorough preprocessing procedure to refine its quality and ensure its relevance to the model's convergence. This preprocessed data was then used for training multiple machine learning models to forecast crop yield. This way, precise predictions regarding the yield of all ten crops can be obtained. This will provide farmers with invaluable insights into their crop rotation strategies. By leveraging these predictions, farmers can optimize their crop production cycles.

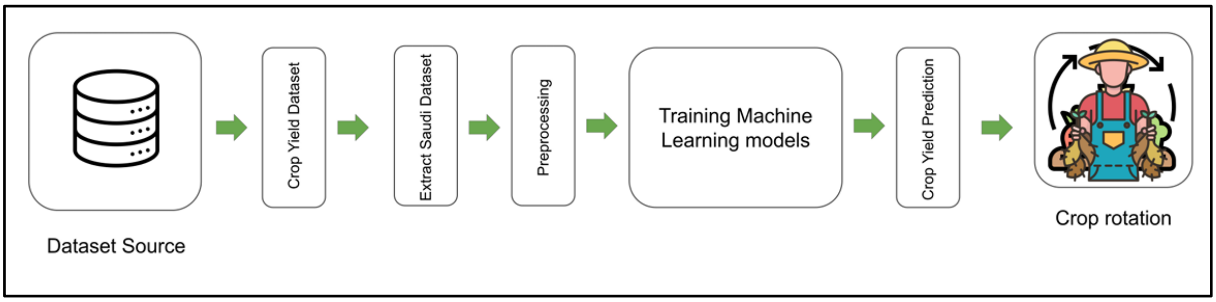


Figure 2. The general scheme outlining the proposed system.

4.1. Data preprocessing

Eighty percent of the crop yield prediction data was used for training, with the remaining portion used for testing. The categorical variables of two columns (Item and Area) in the dataset were initially encoded using one-hot encoding. Categorical data consists of label values rather than numerical values. In many cases, the number of potential values is limited to a specific set, such as the values for objects and nations. Many machine learning and deep learning algorithms cannot directly process labeled data, requiring all input and output variables to be numerical. Hence, categorical data were translated into a numerical format. One-hot encoding facilitates the conversion of categorical data into a form usable by machine learning and deep learning algorithms, thereby enhancing prediction accuracy.

The dataset contains features with widely varying magnitudes, units, and ranges. Features with high magnitudes exert a greater influence on distance estimates compared to those with low magnitudes. To mitigate this effect, it was necessary to standardize all characteristics to the same magnitude, a process achieved through feature scaling. As the dataset has a small standard deviation, the Min-Max scaling algorithm was chosen for scaling. The scaling algorithm transformed the features into a range of 0 to 1. This process involves subtracting the mean of the column from each value and then dividing by the range:

$$\frac{x_i - \min(x)}{\max(x) - \min(x)} \tag{1}$$

4.2. Crop yield prediction

Several machine learning algorithms were trained on the crop yield prediction dataset. These include random forest (RF), K-nearest neighbor (KNN), XGBoost ensemble, decision tree (DT), and a Bagging regressor ensemble meta-estimator. The models were trained and tested using optimal hyperparameter values. Fine-tuning was conducted throughout the manual search process, selecting various combinations of hyperparameter values to assess the models' performance during training. The hyperparameters tuned included the number of estimators for the Bagging regressor, random forest regressor, and XGBoost ensemble model, the max-depth value for the decision tree regressor, and the K-value for KNN. The peak performance achieved through fine-tuning is illustrated in Figures 4–6. Jupyter Notebook version 7.2 was used to train the models, utilizing Python version 3.12.0 and an NVIDIA T4 GPU.

4.2.1. Random forest regressor

A random forest is a meta-estimator that uses averaging to increase prediction accuracy and manage over-fitting. It fits several decision tree regressors on different subsamples of the dataset. The optimal split method is employed by forest trees. Individually, every decision tree shows a considerable variance. However, the variance decreases by combining every decision tree simultaneously. This is due to the fact that every decision tree is meticulously trained on a particular sample of data, guaranteeing that the result is reliant on the combined predictions of several decision trees rather than just one.

Figure 3 illustrates the process where the primary training dataset is split into n samples using bootstrapping. A random forest of decision trees T_k was grown to the bootstrapped data by recursively repeating the following processes for each terminal node until the minimal node size was achieved:

1. Select r_i variables arbitrarily from the R-set of training variables.
2. Select the optimal variable/split point from the r_i.
3. Divide the node into two daughter nodes.

The mean of the results from each individual decision tree is calculated to obtain the final output for regression tasks. This process is commonly referred to as aggregation. Following aggregation, a final prediction is generated:

$$\frac{1}{k}\sum_{i=1}^{k} T_i(x) \tag{2}$$

4.2.2. K-nearest neighbors regression

K-nearest neighbors (KNN) is a non-parametric machine learning method. For numerical regression, predictions are made by locating the K closest data points to an input value and averaging their target values. First, a hyperparameter K value is selected. K is the number of nearby neighbors that must be taken into account during prediction. Next, the Euclidean distance (Eq. 3) is used to calculate how similar the target and training data points are to each other.

$$d(x, X_i) = \sqrt{\sum_{k=1}^{d}(x_k - X_{ik})^2} \tag{3}$$

Here, the training dataset X contains n data points. Each n is represented as a d-dimensional vector of features x_i, with Y denoting the label for each x. Every data point in the dataset has its distance from the target point estimated. The closest neighbors are the k data points that have the least distances to the target position. The projected value for the new data point is calculated as the means of the desired values of the K nearest neighbors. This can be expressed as an arithmetic mean. The method computes the weighted mean of y of the K nearest neighbors and applies it to the projected value for x.

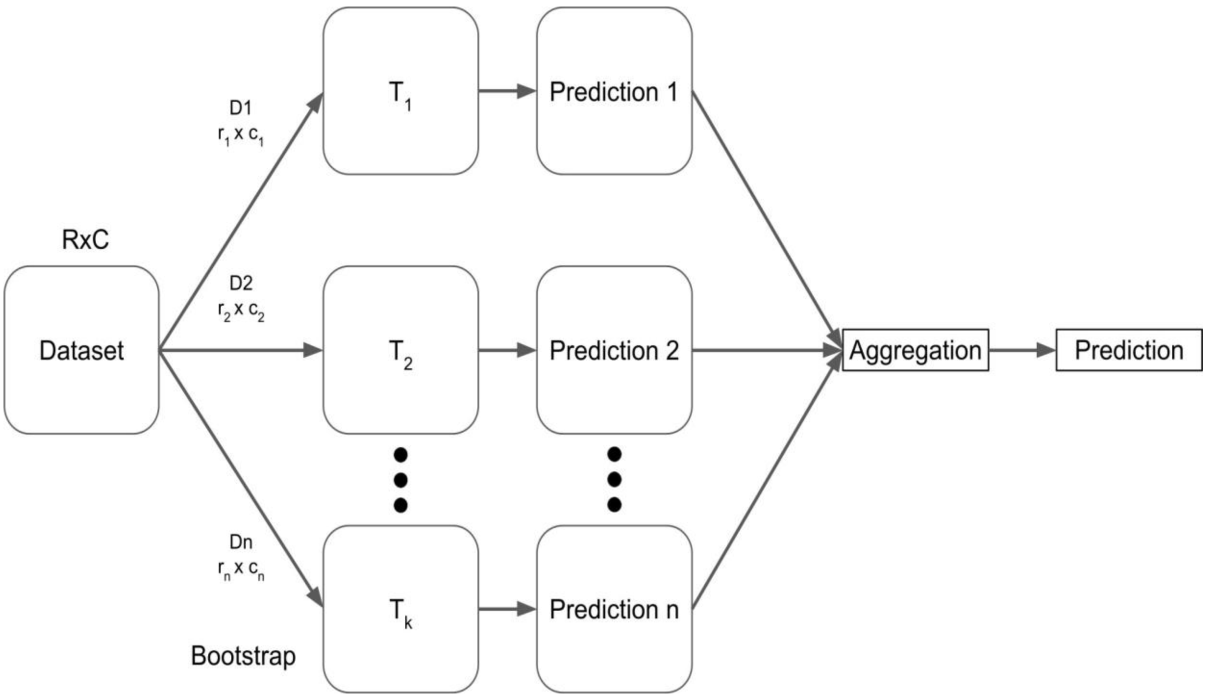


Figure 3. Random forest regressor architecture.

4.2.3. XGBoost ensemble

Extreme gradient boosting (XGBoost) is a highly regarded machine-learning technique renowned for its efficiency, speed, and precision. It is classified within the boosting algorithm family as an ensemble learning methodology that combines the predictions of numerous weak learners. It adopts a boosting strategy to create an exceptionally accurate ensemble model, with each weak learner tasked with correcting the errors made by its predecessors. The gradient optimization approach aims to optimize a cost function by adjusting the model's parameters based on erroneous gradients. This technique introduces the concept of "gradient boosting with decision trees", wherein the objective function is systematically minimized by evaluating the importance of each decision tree added to the ensemble. Additionally, by incorporating a regularization term and employing a more sophisticated optimization technique, XGBoost enhances both accuracy and efficiency.

4.2.4. Decision tree

A decision tree is a nonparametric supervised learning technique employed in both classification and regression tasks within the field of machine learning. This method is distinguished by its hierarchical tree structure, which encompasses essential elements such as a root node, branches, internal nodes, and leaf nodes. There are usually several types of decision trees, i.e., ID3, C4.5, and CART. This work is based on CART, which uses Gini impurity to choose which attribute to split on. Gini impurity indicates how frequently a random machine learning selected property is misclassified. It creates binary splits, and a lower value of Gini impurity is ideal.

The decision tree begins working from the root node *S* with the complete training dataset. It then determines the optimal attribute in the dataset either employing the information gain or the Gini index criterion. Subsequently, S is partitioned into subsets ($s_1, s_2, \ldots, s_i$) containing feasible values for the identified best attributes. Afterward, a decision tree node is generated to encapsulate the optimal attribute. The process recurs as new decision trees are constructed recursively using the subsets ($s_1, s_2, \ldots, s_i$) of the previously derived dataset. This iterative process persists until further classification of the nodes is unattainable, culminating in the identification of leaf nodes.

4.2.5. Bagging regressor ensemble meta-estimator

A bagging regressor serves as a meta-estimator that employs base regressors on random subsets of the original dataset, subsequently aggregating their individual predictions through voting or averaging to produce a final forecast. This meta-estimator is frequently utilized to mitigate the variance of a black-box estimator by introducing randomization into its construction process, ultimately forming an ensemble model.

5. Results and discussion

The trained models were evaluated using several performance metrics, including R^2 score, RMSE, and MAE. To define the metrics, we use Equations 4–6:

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y}_i)^2} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \quad (5)$$

$$MAE = \frac{|(y_i - \hat{y}_i)|}{N} \quad (6)$$

Here, y_i represents the actual value, $\hat{y}_i$ represents the predicted value, $\bar{y}_i$ represents the mean value of the i-th data point, and N is the total number of data points.

The ablation study depicted in Figure 4 presents various hyperparameter tunings for different machine learning models in terms of their R^2 scores. The R^2 scores measure the proportion of the variance in the crop yield that is predictable from the input features. Higher R^2 scores indicate better predictive performance. As observed, in Figure 4(a), increasing the number of estimators did not lead to an improvement in the performance of the Bagging regressor. Notably, the optimal number of estimators ranging between 20 and 40 yielded an R^2 score of 0.9738. Figure 4(c) illustrates that the R^2 score for KNN reached its peak at a K value of 8, achieving a score of 0.9573, while deviations from this optimal K value led to a decline in performance. Additionally, Figure 4(b) highlights that an optimal decision tree depth value of 20 resulted in an R^2 score of 0.9628, with performance saturation observed beyond this value. Similarly, Figure 4(d) and (e) shows that the optimal values of the number of estimators are 140 and 320, with R^2 scores of 0.9740 and 0.9745 for random forest and XGBoost ensemble models, respectively.

Figure 5 exhibits the comparison between actual and predicted values produced by the machine learning models employed in this study, optimized with optimal hyperparameters. The closer the points are to the blue straight line with a slope of 1, the more accurate the predictions are. Table 3 shows the R^2 scores, root mean square error (RMSE), and mean absolute error (MAE) with optimal hyperparameter tuning of the different machine learning models used in this work for the crop yield prediction test dataset. Previously, [51] worked on this dataset and achieved a R^2 score of 0.9570. Our KNN model could not surpass their score, but the decision tree regressor (DT) trained in this work slightly outperformed the model developed by [51] with an R^2 score of 0.9628. The model had an RMSE value of 16697.12 and an MAE of 3751.65. The Bagging regressor performed even better, with an R^2 score of 0.9738, RMSE of 14010.27, and MAE of 3294.23. The random forest regressor (RF) had a slightly better performance with an R^2 score of 0.9740, RMSE of 13781.94, and MAE of 3076.19. The XGBoost ensemble clearly outperformed all other models with an R^2 score of 0.9745, RMSE of 15803.15, and MAE of 2681.33. Finally, Figure 6 presents the progression curve of number of estimators hyperparameter value and accuracy (R^2 score) in XGBoost ensemble.

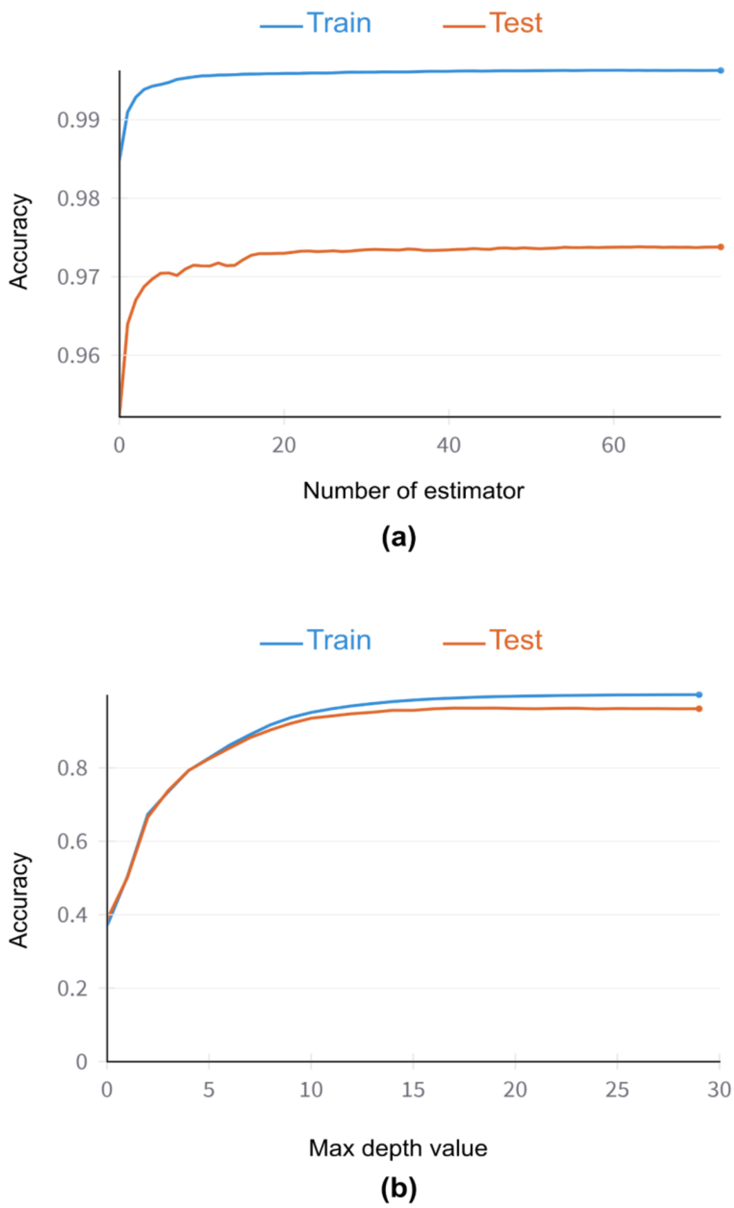
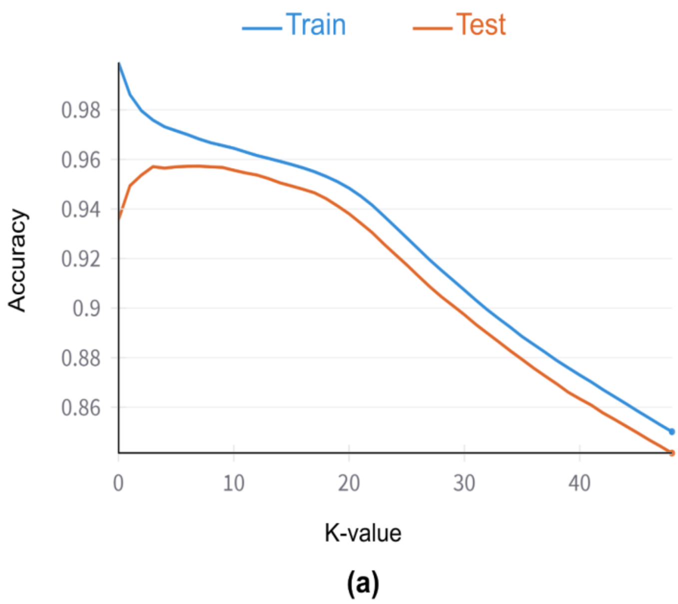


Figure 4. Progression curve of (a) number of estimators hyperparameter value vs. accuracy (R^2 score) in Bagging regressor; (b) Max depth hyperparameter value of decision tree regressor vs. accuracy (R^2 score).



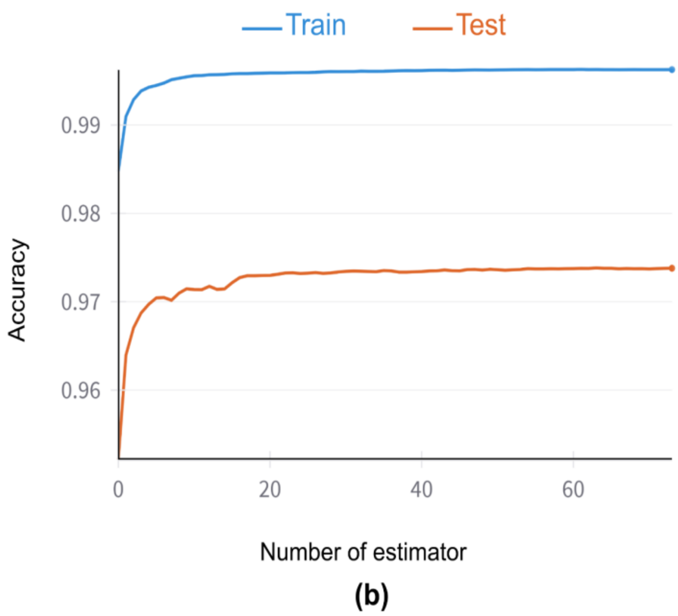


Figure 5. Progression curve of (a) K-value vs. accuracy (R^2 score) in KNN; (b) number of estimators hyperparameter value vs. accuracy (R^2 score) in random forest regressor.

Table 3. R^2 score, RMSE, and MAE achieved using machine learning models for the crop yield prediction test dataset.

Model	R^2 score	RMSE	MAE
KNN	0.9573	17603.76	4650.21
AdaBoost regressor with decision tree [51]	0.9570	-	-
Decision tree regressor (DT)	0.9628	16697.12	3751.65
Bagging regressor	0.9738	14010.27	3294.23
Random forest regressor (RF)	0.9740	13781.94	3076.19
XGBoost ensemble	0.9745	15803.15	2681.33

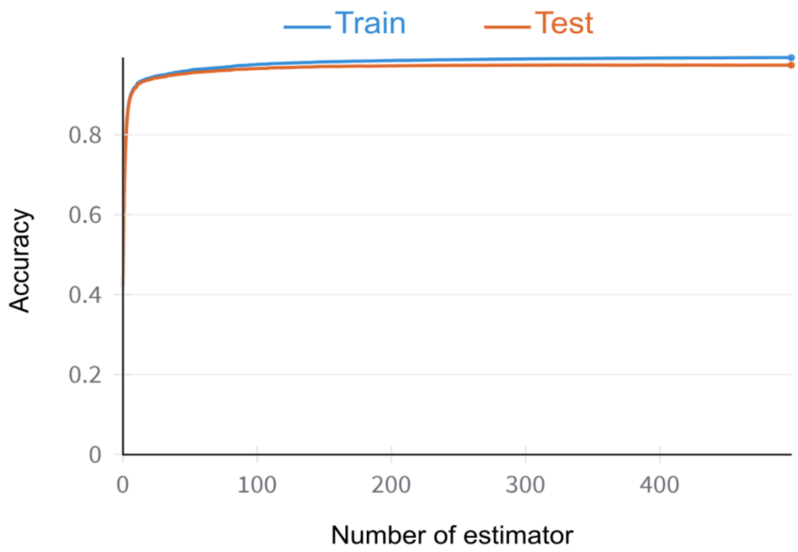


Figure 6. Progression curve of number of estimators hyperparameter value vs. accuracy (R^2 score) in XGBoost ensemble.

The performance of various machine learning models in predicting crop yields across different crop types was also evaluated. The models trained on the crop yield prediction training dataset were tested for every crop in the test dataset. The results presented in Table 4 display the prediction R^2 scores obtained for each crop using different models. Each row within the table corresponds to a specific crop, while each column represents a machine-learning model employed for prediction. Analysis of the results reveals that certain models exhibit superior performance for specific crops. For instance, the Bagging regressor, random forest regressor, and XGBoost ensemble consistently achieve high R^2 scores across multiple crops, showcasing their effectiveness in yield prediction. Conversely, for crops such as rice/paddy and sweet potatoes, the XGBoost ensemble model demonstrates superior performance compared with other models, suggesting its suitability for these particular crops. These findings offer valuable insights into the selection of appropriate machine learning models for crop yield prediction, thereby facilitating the optimization of agricultural practices and resource allocation.

Table 4. Prediction R^2 scores of each crop for different models.

Crop	KNN	DT regressor	Bagging regressor	RF regressor	XGBoost ensemble
Maize	0.9686	0.9378	0.9815	0.9141	0.986
Potato	0.9778	0.9227	0.9940	0.9421	0.994
Rice/paddy	0.9437	0.9452	0.9470	0.9482	0.9737
Sorghum	0.9539	0.9912	0.9950	0.9970	0.964
Soybean	0.9559	0.9627	0.9377	0.9956	0.9668
Wheat	0.9477	0.9845	0.9750	0.9831	0.9652
Cassava	0.9790	0.9660	0.9557	0.9746	0.9459
Sweet potato	0.9478	0.9904	0.9576	0.9902	0.9999
Plantain	0.9406	0.9820	0.9960	0.9999	0.968
Yam	0.9580	0.9431	0.9998	0.9989	0.9733

Hence, by utilizing the crop yield predictions from the best-trained model, the XGBoost ensemble, farmers can anticipate the upcoming season's crop yield. Armed with this information, farmers can determine which crops are likely to have the highest production in the upcoming season. Consequently, they can rotate their crops without the need for tilling.

While the dataset used in this study lacks specific date crop data, it does encompass some of the most widely consumed crops globally. Notably, a few of these crops align with the top-grown crops in MENA nations. Consequently, the findings from this study hold promise for date crop prediction as well, utilizing transfer learning techniques and a tailored dataset that includes date-related data.

Figure 7 compares the actual crop yield values with the predicted values using two machine learning models: the Random Forest Regressor (Figure 7a) and the XGBoost Ensemble (Figure 7b). The two models reveal the predictions closely align with the actual yields as it shows a high degree of accuracy indicating by the closeness of points to the diagonal line.

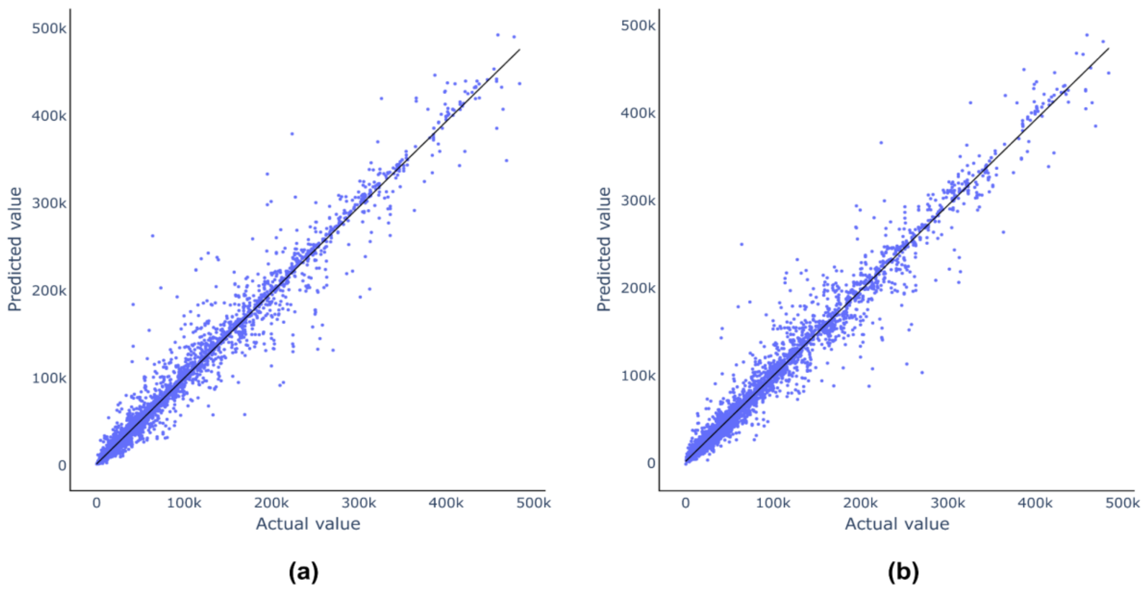


Figure 7. Actual vs. predicted data using (a) random forest regressor and (b) XGBoost ensemble.

Figure 8 indicates the comparison between actual and predicted crop yield values for two additional machine learning models such as K-Nearest Neighbors (KNN) as shown in figure 8a and the Decision Tree Regressor as shown in figure 8b. The performance of these models shows some variance from the actual values but is still relatively accurate, with some divergence from the diagonal line, indicating slightly lower prediction accuracy compared to the Random Forest and XGBoost models.

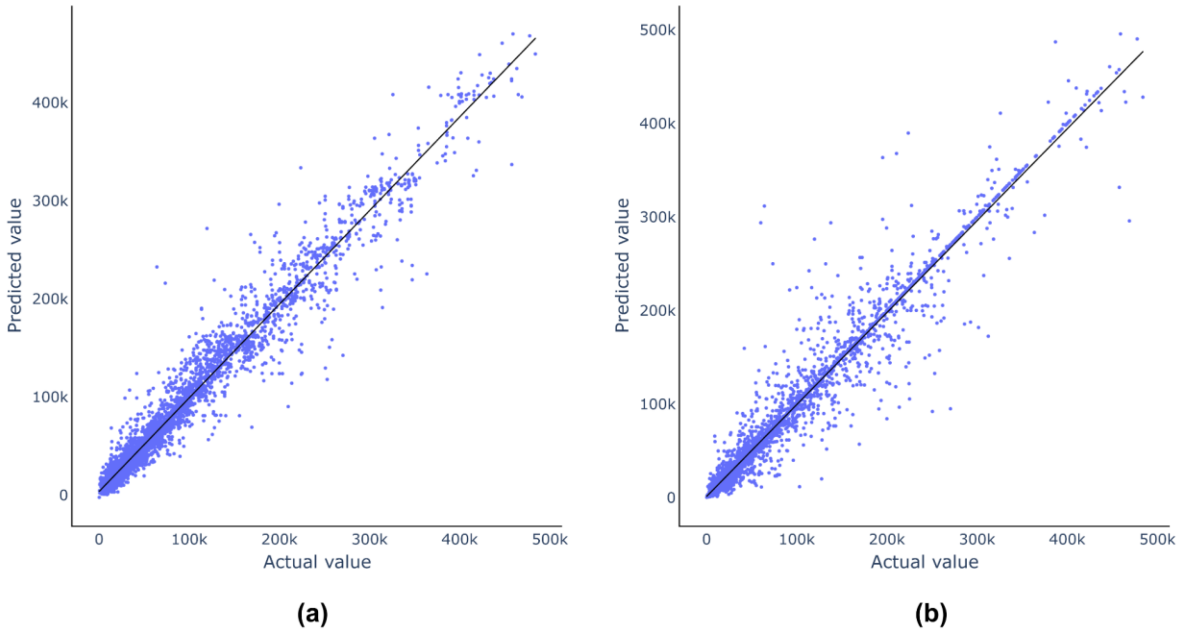


Figure 8. Actual vs. predicted data using (a) KNN and (b) decision tree regressor.

Figure 9 shows the actual versus predicted crop yield values using the Bagging Regressor model. The points are close to the diagonal line, indicating that the Bagging Regressor is also highly effective in predicting crop yields, similar to the Random Forest and XGBoost models, though its performance might be slightly less optimal compared to XGBoost.

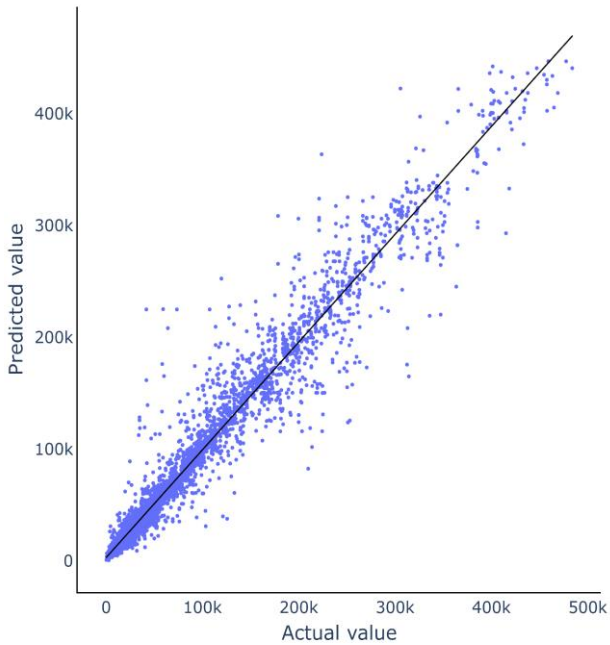


Figure 9. Actual vs. predicted data using Bagging regressor.

6. Theoretical and practical implications

The scientific community can benefit from this study in several ways. First, the research offers methodological insights into the development and training of machine learning models for crop yield prediction, providing a framework for similar research endeavors. By understanding how machine learning-based systems can optimize crop rotation and mitigate climate change impacts, researchers can identify practical applications for their work, tailored to different agricultural contexts. The collaborative nature of the research underscores the importance of interdisciplinary collaboration, providing researchers with opportunities to engage with stakeholders from various sectors to co-create knowledge and develop innovative solutions to agricultural challenges.

In practice, this study offers valuable insights for policymakers, agricultural practitioners, and stakeholders in Saudi Arabia. By leveraging machine learning models for prediction, farmers can make informed decisions about crop rotation without tilling their land, thereby reducing carbon emissions. Additionally, the integration of sustainable finance mechanisms can support this initiative by providing financial incentives for farmers who adopt eco-friendly practices. Specifically, microfinancing options or green loans can be made available to farmers who invest in sustainable technologies, such as those developed in this study. These financial products could lower the barrier to entry for adopting innovative agricultural practices that enhance productivity while also promoting environmental stewardship. Furthermore, a mobile application can be developed containing the best model to facilitate easy access to the system for farmers. Overall, the implementation of the proposed system and aligning it with sustainable finance strategies has the potential to significantly improve crop production and soil health and contribute to the broader goal of climate change mitigation in Saudi Arabia.

7. Conclusion

This study proposed a novel strategy for crop rotation without tilling through crop yield prediction and machine learning to address the pressing need for proactive measures to mitigate the challenges of climate change in agriculture. By evaluating the performance of various machine learning algorithms and combining multiple data sources, the study achieved improved accuracy. Optimization of hyperparameters and assessment of crop-wise prediction performance for random forest, K-nearest neighbor, XGBoost ensemble, decision tree, and a Bagging regressor ensemble meta-estimator demonstrated the effectiveness of machine learning in informing crop rotation strategies and resource management decisions, offering practical solutions for climate resilience and sustainable agriculture in Saudi Arabia. This underscores the promising role of machine learning in improving crop yield forecasts and guiding sustainable agricultural methods. Leveraging the machine learning models trained on the crop yield prediction dataset, stakeholders can gain valuable insights into optimal crop rotation strategies and resource management techniques. Collaboration between farmers, researchers, and policymakers is paramount in translating these insights into actionable strategies for climate resilience and sustainable agriculture. As the global community continues to grapple with the impacts of climate change, initiatives like these offer hope for a more resilient and sustainable agricultural future.

The dataset used in the study does not include data on date fruit, a significant crop in Saudi Arabia. This limitation hinders the applicability of the findings to date production, a critical agricultural output in the region. While the dataset includes some widely consumed crops, it does not cover all crops grown in Saudi Arabia, restricting the study's scope in providing comprehensive crop yield predictions for the region.

In the future, a benchmark dataset tailored for Saudi Arabia could be developed. This dataset would offer deeper insights, aiding in the creation of a precise crop yield prediction system. By incorporating a broader range of climatic and soil factors, this system would facilitate more effective crop rotation strategies, crucial for mitigating the impacts of climate change and ensuring sustainable agricultural practices.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This project was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, under Grant No. IFPIP: 618-849-1443. The authors, therefore, acknowledge with thanks the DSR for financial support.

Conflict of interest

The authors declare no conflict of interest.

Authors Contributions

Conceptualization: M.M.I. and M.A.; data curation: A.K.M.M.; formal analysis: R.I. and A.K.M.M.; investigation: M.M.I. and R.S.A.; methodology: R.I. and A.K.M.M.; project administration: M.A. and R.S.A.; resources: M.A. and R.S.A.; software: R.I.; supervision: M.A. and R.S.A.; validation: M.M.I.; visualization: M.M.I. and A.K.M.M.; writing-original draft: all authors; writing-review and editing: all authors. Finally, all authors have read and agreed to the published version of the manuscript.

References

1. Romm JJ (2022) *Climate change: What everyone needs to know.* Oxford University Press.
2. DeNicola E, Aburizaiza OS, Siddique A, et al. (2015) Climate change and water scarcity: The case of Saudi Arabia. *Ann Global Health* 81: 342–353. https://doi.org/10.1016/j.aogh.2015.08.005
3. Field CB (2012) *Managing the risks of extreme events and disasters to advance climate change adaptation: Special report of the intergovernmental panel on climate change.* Cambridge University Press. https://doi.org/10.1017/CBO9781139177245
4. Masson-Delmotte V, Zhai P, Pörtner HO, et al. (2018) *Global warming of 1.5 ℃: IPCC special report on impacts of global warming of 1.5 ℃ above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty.* Cambridge University Press.

5. Edenhofer O (2015) *Climate change 2014: Mitigation of climate change.* Cambridge University Press.
6. Montoya JH, Tsai C, Vojvodic A, et al. (2015) The challenge of electro- chemical ammonia synthesis: A new perspective on the role of nitrogen scaling relations. *ChemSusChem* 8: 2180–2186. https://doi.org/10.1002/cssc.201500322
7. Robertson GP, Vitousek PM (2009) Nitrogen in agriculture: Balancing the cost of an essential resource. *Ann Rev Environ Resour* 34: 97–125. https://doi.org/10.1146/annurev.environ.032108.105046
8. Hawken P (2017) *Drawdown: The most comprehensive plan ever proposed to reverse global warming.* Penguin.
9. Wahabzada M, Mahlein AK, Bauckhage C, et al. (2016) Plant phenotyping using probabilistic topic models: uncovering the hyperspectral language of plants. *Sci Rep* 6: 22482. https://doi.org/10.1038/srep22482
10. Liakos KG, Busato P, Moshou D, et al. (2018) Machine learning in agriculture: A review. *Sensors* 18: 2674. https://doi.org/10.3390/s18082674
11. Rossel RAV, Bouma J (2016) Soil sensing: A new paradigm for agriculture. *Agric Syst* 148: 71–74. https://doi.org/10.1016/j.agsy.2016.07.001
12. You J, Li X, Low M, et al. (2017) Deep gaussian process for crop yield prediction based on remote sensing data. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 31: No. 1. https://doi.org/10.1609/aaai.v31i1.11172
13. Ma W, Nowocin K, Marathe N, et al. (2019) An interpretable produce price forecasting system for small and marginal farmers in india using collaborative fil- tering and adaptive nearest neighbors. In: *ICTD '19: Proceedings of the Tenth International Conference on Information and Communication Technologies and Development,* Association for Computing Machinery, New York, NY, USA, Article 6, 1–11. https://doi.org/10.1145/3287098.3287100
14. Alskaf K, Mooney S, Sparkes D, et al. (2021) Short-term impacts of different tillage practices and plant residue retention on soil physical properties and greenhouse gas emissions. *Soil Tillage Res* 206: 104803. https://doi.org/10.1016/j.still.2020.104803
15. Siwar C, Alam MM, Murad MW, et al. (2009) A review of the linkages between climate change, agricultural sustainability and poverty in Malaysia. *Int Rev Bus Res Pap* 5: 309–321.
16. Kurukulasuriya P, Rosenthal S (2003) Climate change and agriculture. *World Bank Environment Department Paper 91.*
17. Core Writing Team, Pachauri RK, Reisinger A (2007) Climate Change 2007: Synthesis Report. IPCC Geneva, Switzerland. Available from: https://www.ipcc.ch/site/assets/uploads/2018/02/ar4_syr_full_report.pdf.
18. IPCC (2014) Climate Change 2014—Impacts, Adaptation, and Vulnerability: Part A: Global and Sectoral Aspects. Cambridge University Press. Available from: http://www.cambridge.org/9781107641655.
19. Webster M, Forest C, Reilly J, et al. (2003) Uncertainty analysis of climate change and policy response. *Clim Change* 61: 295–320. https://doi.org/10.1023/B:CLIM.0000004564.09961.9f
20. Zwiers FW (2002) The 20-year forecast. *Nature* 416: 690–691. https://doi.org/10.1038/416690a
21. Stern N (2007) *The economics of climate change: The stern review.* Cambridge University Press.
22. Emissions gap report 2022 (2022) United Nations Environment Programme.
23. Alkolibi FM (2002) Possible effects of global warming on agriculture and water resources in Saudi Arabia: Impacts and responses. *Clim Change* 54: 225–245. https://doi.org/10.1023/A:1015777403153

24. Aleid SM, Al-Khayri JM, Al-Bahrany AM (2015) Date palm status and perspective in Saudi Arabia. In: Al-Khayri JM, Al-Khayri SM, Johnson DV (Eds.), *Date Palm Genetic Resources and Utilization Volume 2: Asia and Europe,* 49–95. https://doi.org/10.1007/978-94-017-9707-8_3
25. Assiri A, Darfaoui E (2009) Response to climate change in the kingdom of Saudi Arabia. A report prepared for FAO-RNE 17.
26. Allbed A, Kumar L, Shabani F (2017) Climate change impacts on date palm cultivation in Saudi Arabia. *J Agric Sci* 155: 1203–1218. https://doi.org/10.1017/S0021859617000260
27. Global Arab Network (2010) Great transition—Saudi Arabia planting new seeds. Available from: https://www.farmlandgrab.org/post/view/12434-great-transition-saudi-arabia-planting-new-seeds.
28. Mbaga MD (2013) Alternative mechanisms for achieving food security in Oman. *Agric Food Sec* 2: Article number: 3, 1–11. https://doi.org/10.1186/2048-7010-2-3
29. Haque MI, Khan MR (2022) Impact of climate change on food security in Saudi Arabia: A roadmap to agriculture-water sustainability. *J Agribus Dev Emerging Econ* 12: No. 1, 1–18. https://doi.org/10.1108/JADEE-06-2020-0127
30. Parry M, Canziani O, Palutikof J, et al. (2007) Climate change 2007: Impacts, Adaptation and Vulnerability. Available from: https://www.ipcc.ch/site/assets/uploads/2018/03/ar4_wg2_full_report.pdf.
31. Baig MB, Straquadine GS (2014) Sustainable agriculture and rural development in the kingdom of Saudi Arabia: Implications for agricultural extension and education. In: Behnassi M, Muteng'e MS, Ramachandran G, et al. (Eds.), *Vulnerability of agriculture, water and fisheries to climate change: Toward sustainable adaptation strategies,* 101–116. https://doi.org/10.1007/978-94-017-8962-2_7
32. Al-Karaki GN, Al-Hashimi M (2012) Green fodder production and water use efficiency of some forage crops under hydroponic conditions. *Int Scholarly Res Not* 2012: 924672. https://doi.org/10.5402/2012/924672
33. Fiaz S, Noor MA, Aldosri FO (2018) Achieving food security in the kingdom of Saudi Arabia through innovation: Potential role of agricultural extension. *J Saudi Soc Agric Sci* 17: 365–375. https://doi.org/10.1016/j.jssas.2016.09.001
34. Frederick KD, Kneese AV (1990) Reallocation by markets and prices. *Clim Change US Water Resour* 1990: 395–419.
35. Nguyen N, Drakou EG (2021) Farmers intention to adopt sustainable agriculture hinges on climate awareness: The case of Vietnamese coffee. *J Cleaner Prod* 303: 126828. https://doi.org/10.1016/j.jclepro.2021.126828
36. Al-Shayaa MS, Baig MB, Straquadine GS (2012) Agricultural extension in the kingdom of Saudi Arabia: Difficult present and demanding future. *J Anim Plant Sci* 22: 239–246.
37. AL-Subaiee SS (2023) Extension agents' perceptions regarding sustainable agriculture in the Riyadh region of Saudi Arabia. Pennsylvania State University.
38. Kassie M, Zikhali P (2009) Brief on sustainable agriculture. In: *Expert Group Meeting on "Sustainable Land Management and Agricultural Practices in Africa: Bridging the Gap Between Research and Farmers",* Gothenburg, Sweden, 16–17.
39. Adger WN (1999) Social vulnerability to climate change and extremes in coastal Vietnam. *World Dev* 27: 249–269.
40. MEWA (2016) National Water Strategy 2030: Towards Sustainable Water Sector that Develops and Conserves Water Resources. Available from: https://www.fao.org/faolex/results/details/en/c/LEX-FAOC191510/.

41. Alotaibi BA, Kassem HS, Abdullah AZ, et al. (2020) Farmers' awareness of agri-environmental legislation in Saudi Arabia. *Land Use Policy* 99: 104902. https://doi.org/10.1016/j.landusepol.2020.104902
42. Alotaibi BA, Kassem HS, Nayak RK, et al. (2020) Farmers' beliefs and concerns about climate change: an assessment from southern Saudi Arabia. *Agriculture* 10: 253. https://doi.org/10.3390/agriculture10070253
43. Alotaibi BA, Abbas A, Ullah R, et al. (2021) Climate change concerns of Saudi Arabian farmers: The drivers and their role in perceived capacity building needs for adaptation. *Sustainability* 13: 12677. https://doi.org/10.3390/su132212677
44. Al-Zaidi A, Elhag E, Al-Otaibi S, et al. (2011) Negative effects of pesticides on the environment and the farmers awareness in Saudi Arabia: A case study. *J Anim Plant Sci* 21: 605–611.
45. Almutawa AA (2022) Date production in the Al-Hassa region, Saudi Arabia in the face of climate change. *J Water Clim Change* 13: 2627–2647. https://doi.org/10.2166/wcc.2022.461
46. Ji B, Sun Y, Yang S, et al. (2007) Artificial neural networks for rice yield prediction in mountainous regions. *J Agric Sci* 145: 249–261. https://doi.org/10.1017/S0021859606006691
47. Drummond ST, Sudduth KA, Joshi A, et al. (2003) Statis- tical and neural methods for site–specific yield prediction. *Trans ASAE* 46: 5–14. https://doi.org/10.13031/2013.12541
48. Kaggle (2024) Crop Yield Prediction Dataset. Available from: https://www.kaggle.com/datasets/patelris/crop-yield-prediction-dataset.
49. Food and Agricultural Organization of the United Nations. Available from: https://www.fao.org/home/en/.
50. World Bank Open Data. Available from: https://data.worldbank.org/.
51. Keerthana M, Meghana K, Pravallika S, et al. (2021) An ensemble algorithm for crop yield prediction. In: *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV),* IEEE, 963–970. https://doi.org/10.1109/ICICV50876.2021.9388479