JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

HARNESSING LARGE LANGUAGE MODELS FOR AGRICULTURAL INNOVATION: OPPORTUNITIES, CHALLENGES, AND FUTURE DIRECTIONS

[1]Husain H. Dawoodi
[1]System Analyst
[1]School of Computer Sciences,
[1]Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon, India

Abstract: The rapid advancements in artificial intelligence, particularly in the domain of large language models (LLMs), have opened new avenues for transforming various sectors, including agriculture. LLMs, such as GPT-4, have demonstrated remarkable capabilities in natural language understanding, text generation, and data analysis, making them valuable tools for enhancing agricultural practices. This review explores the diverse applications of LLMs in agriculture, ranging from precision farming, crop disease detection, and weather forecasting to supply chain management, policy recommendations, and farmer advisory systems. By leveraging vast datasets and complex models, LLMs can assist in optimizing resource allocation, improving crop yields, and providing real-time decision support. The paper also highlights the challenges and limitations of implementing these models in rural settings, particularly in terms of data availability, accessibility, and computational infrastructure. This review aims to provide insights into the current state of LLM applications in agriculture and suggests future research directions to harness their full potential for sustainable agricultural development.

***Index Terms* – Large Language Models, Applications in Agriculture.**

I. INTRODUCTION

The agricultural sector faces critical challenges in the 21st century, from population growth and climate change to resource scarcity and evolving market demands. As traditional agricultural practices struggle to keep up with these dynamics, the need for technological solutions is more pressing than ever. In recent years, artificial intelligence (AI) has emerged as a powerful tool to drive innovation in agriculture, improving productivity and sustainability. Among AI's various subsets, Large Language Models (LLMs), a category of deep learning models trained on vast amounts of text data, are gaining prominence due to their ability to understand and generate human-like language. These models, such as OpenAI's GPT-4 and Google's BERT, have revolutionized numerous fields, including healthcare, finance, and customer service. However, their potential in agriculture is still under-explored.

Large Language Models (LLMs) are a type of AI model that rely on transformers, a neural network architecture that excels at processing sequences of data, particularly text. LLMs are trained on vast datasets, which include diverse sources like books, websites, research articles, and social media, allowing them to understand and generate coherent and contextually relevant language. This capability enables them to perform a range of tasks such as natural language processing (NLP), machine translation, data analysis, and decision support systems. In agriculture, these models could potentially transform areas like precision farming, crop disease detection, weather forecasting, supply chain management, and even farmer advisory services.

Despite the immense promise, the application of LLMs in agriculture is still in its infancy. Most current AI applications in agriculture focus on computer vision for crop monitoring, machine learning models for yield prediction, and sensor-based IoT systems for precision farming. LLMs, which have demonstrated success in textual analysis and decision-making in other sectors, are only beginning to be considered for agricultural applications. The gap between the advancements in LLMs and their practical utility in agriculture stems from several challenges and limitations.

The current limitations of Large Language Models (LLMs) in agriculture stem from various technical, infrastructural, and data-related challenges. One of the primary limitations is the lack of domain-specific datasets; agricultural data is often fragmented, region-specific, and not standardized, making it difficult for LLMs to generate precise, contextually relevant outputs. Additionally, high computational requirements hinder the deployment of these models in rural areas, where access to robust computing infrastructure and stable internet connections is limited. Another challenge is the multilingual gap, as most LLMs are trained on widely spoken languages, limiting their effectiveness in rural regions where farmers speak local dialects. Moreover, LLMs struggle with real-time decision-making, which is crucial in agriculture for responding to dynamic environmental changes. Finally, ethical concerns arise regarding data bias, misinformation, and the risk of LLMs providing inaccurate recommendations that could adversely impact farming outcomes (Tzachor et al., 2023; Wang et al., 2023). While these limitations are significant, there is growing recognition of the potential for LLMs to fill critical gaps in agricultural research and practice. The existing literature predominantly addresses the applications of machine learning and AI in agriculture through the lens of structured data, such as satellite imagery, sensor data, or statistical crop models. However, the application of LLMs—which excel in processing unstructured text data, analyzing large-scale knowledge, and generating context-sensitive advice—has not yet been systematically explored. Furthermore, current research tends to overlook how LLMs can address challenges unique to agriculture, such as multilingual farmer support, policy recommendations, and real-time decision-making.

This paper aims to fill this research gap by providing a comprehensive review of how LLMs can be applied in agriculture, detailing their potential to enhance various farming practices, and outlining the challenges that need to be overcome for successful implementation. Through this review, we will explore how LLMs can aid in precision farming, crop monitoring, disease detection, and forecasting by processing large-scale text and knowledge-based data. Examine the role of LLMs in improving multilingual advisory systems, offering real-time decision support to farmers in multiple languages, particularly those in rural and underserved regions. Investigate the integration of LLMs with IoT systems and other sensor-based technologies to facilitate real-time data analysis and decision-making in agriculture. And discuss the ethical concerns and potential biases associated with the application of LLMs in agriculture and propose strategies to mitigate these risks.

In doing so, this paper seeks to bridge the gap between the theoretical potential of LLMs and their practical applications in agriculture, offering new insights into how these advanced AI models can contribute to solving the pressing challenges of modern agriculture

II. LITERATURE REVIEW

Disease Detection and Diagnosis: Recent studies highlight the use of LLMs combined with deep learning and agricultural knowledge graphs to enhance the detection and diagnosis of plant diseases. For example, a system developed to detect *Elaeagnus angustifolia* disease integrates LLMs, knowledge graphs, and graph neural networks. This approach has significantly improved the accuracy, recall, and precision of disease detection over traditional methods, demonstrating the potential of LLMs in optimizing plant protection strategies (Tzachor et al., 2023; Zhao et al., 2024).

Automated Agricultural Advisory Services: LLMs like GPT-4 are being explored for their potential to assist in agricultural advisory services. These models can provide farmers with timely advice on crop management, pest control, and disease prevention based on real-time data inputs. The ability of LLMs to process large datasets and respond to queries in natural language has made them valuable in rural areas where agricultural extension services are limited (Tzachor et al., 2023).

Smart Agriculture and Data Management: LLMs are increasingly applied to process unstructured agricultural data, such as reports, weather patterns, and soil health data, to provide actionable insights. Embedding-based retrieval models have been developed to enhance the extraction of relevant information from large datasets. This capability is crucial for precision agriculture, where decisions depend on analyzing vast amounts of data from sensors, satellites, and other IoT devices (Wang et al., 2023).

Knowledge Transfer and Education: LLMs are also being utilized for training and educating farmers. By processing technical documents and providing easy-to-understand advice, these models can bridge knowledge gaps, particularly in developing countries. For example, models like ChatAgri have been employed to help classify agricultural texts across different languages, assisting multilingual agricultural extension services (Wang et al., 2023).

III. APPLICATIONS OF LARGE LANGUAGE MODELS IN AGRICULTURE (LLMs)

i. Precision Farming:

Precision farming, a data-driven approach to optimize crop production, has seen significant advancements with the integration of Large Language Models (LLMs). These models are capable of analyzing vast datasets collected from sensors, satellites, and IoT devices to provide real-time insights into agricultural practices such as irrigation, fertilization, and pest control. LLMs, like GPT-4, can process complex, unstructured data such as weather reports, soil health records, and historical crop performance, generating actionable recommendations for farmers to enhance crop yields and resource use efficiency.

Recent research emphasizes the role of LLMs in improving precision farming outcomes. For instance, a study by Peng et al. (2023) demonstrated the use of embedding-based retrieval models in agriculture to extract and process information from diverse datasets, enabling precise decision-making. These models improve the accuracy of predictions by analyzing environmental and crop-specific data, contributing to better water and fertilizer management. Similarly, in the context of plant health monitoring, an LLM-based system integrated with knowledge graphs and neural networks significantly improved the precision of crop disease detection and pest management, reducing over-application of pesticides and optimizing resource use (Tzachor et al., 2023).

Moreover, LLMs' ability to understand and generate human-like language allows them to be integrated with farmer advisory systems, providing farmers with real-time, personalized recommendations. This is particularly beneficial in precision agriculture, where every decision—from planting to harvesting—can have a profound impact on productivity and sustainability. Silva et al. (2023) explored the use of GPT-4 for answering agricultural exams and providing agronomist-level advice, highlighting how such models could support precision farming through better information dissemination and farmer education.

Despite these advancements, there remain challenges, including the need for localized data to fine-tune LLMs for specific regions and crops, as well as high computational costs associated with processing large volumes of real-time data in resource-limited settings. Addressing these challenges is crucial for maximizing the benefits of LLMs in precision farming.

ii. Crop Disease Detection and Monitoring:

Crop disease detection and monitoring are critical components of sustainable agriculture, and the integration of Large Language Models (LLMs) into this field is transforming how farmers and researchers address plant health. LLMs can process large amounts of unstructured data from agricultural reports, satellite imagery, research papers, and farmer feedback to identify patterns and predict the onset of diseases, thereby enabling timely intervention.

Recent research has shown that LLMs can enhance crop disease detection systems by integrating natural language processing (NLP) with machine learning techniques and knowledge graphs. For instance, Rezayi et al. (2022) developed an agricultural language model, AgriBERT, that leverages domain-specific data for matching food and nutrition requirements, which can be adapted for plant health monitoring as well. In addition, Yang et al. (2024) introduced PLLaMa, an open-source language model for plant science, which demonstrates how specialized models can be fine-tuned to detect specific plant diseases and monitor crop health across different regions.

One notable advancement is the integration of LLMs with computer vision technologies, where models can process text data describing visual symptoms of diseases (such as leaf spots or discoloration) alongside actual image data. This fusion enables systems to provide real-time, text-based advice on crop diseases to farmers. A study by Zhao et al. (2023) applied ChatGPT to cross-linguistic agricultural text classification, highlighting its potential in classifying disease reports from various languages, which can be critical in global crop disease monitoring efforts.

In addition to detection, LLMs help generate actionable recommendations for disease prevention. They process not only environmental data but also historical disease outbreaks, enabling predictive monitoring. Wang et al. (2023) explored the development of intelligent question-answering systems based on LLMs for fruit and vegetable disease diagnosis, showing how models can assist farmers by answering disease-related questions based on past experiences and global agricultural knowledge.

However, several challenges remain. The availability of high-quality training data is limited, especially in developing countries, where digitized agricultural records are scarce. Moreover, LLMs often require significant computational power to process large datasets and integrate with real-time sensor networks. Addressing these challenges will require further collaboration between AI developers and agricultural scientists to create more accessible, domain-specific models.

iii. *Weather Forecasting and Climate Adaptation:*

Weather forecasting and climate adaptation are pivotal for modern agriculture, where precise information about weather patterns can significantly impact crop yields, irrigation scheduling, and disaster preparedness. Large Language Models (LLMs) are playing an increasingly important role in processing complex climate data and generating accurate, localized forecasts that help farmers adapt to changing weather conditions and long-term climate shifts.

LLMs can analyze and interpret vast datasets, such as historical weather patterns, satellite data, and real-time sensor inputs, enabling more nuanced weather forecasts. They also excel in synthesizing unstructured data from reports and research papers, providing actionable insights to farmers. For example, Peng et al. (2023) demonstrated how embedding-based retrieval models integrated with LLMs can extract critical climate information from a wide array of unstructured sources, improving the precision of weather forecasts for agriculture. Such models help farmers make informed decisions regarding irrigation, planting, and harvesting based on localized, real-time forecasts.

Moreover, LLMs are aiding in climate adaptation strategies, helping farmers anticipate the effects of long-term climate change. These models can analyze regional climate data, predict future trends, and offer crop-specific advice to mitigate the impacts of extreme weather conditions like droughts, floods, and heatwaves. For example, ChatGPT has been explored in agricultural applications to provide tailored advice based on historical weather data, enabling farmers to adapt their practices to minimize losses from changing climates.

LLMs can also play a role in disaster preparedness by providing early warnings of severe weather events, such as hurricanes or unseasonal rains, which can damage crops. This helps farmers take timely actions like harvesting early, deploying protective measures, or adjusting planting schedules. Studies like Wang et al. (2023) highlight how intelligent question-answering systems powered by LLMs can assist farmers by responding to weather-related queries and offering actionable recommendations based on historical and real-time data.

However, while LLMs offer potential for improved weather forecasting and climate adaptation, challenges persist. The quality of forecasts depends on the availability of accurate and localized climate data, which can be difficult to obtain in some regions. Additionally, the computational demands of processing large datasets and integrating real-time updates require significant infrastructure, which can be a barrier in rural and developing areas. Overcoming these obstacles is essential to fully leverage LLMs for effective climate adaptation in agriculture.

iv. *Farmer Advisory Systems*

Farmer advisory systems are becoming crucial in providing timely, location-specific information to farmers, and Large Language Models (LLMs) are emerging as powerful tools in transforming these systems. LLMs, with their ability to process vast amounts of agricultural data and deliver human-like responses, are proving instrumental in enhancing how advisory services are delivered to farmers, particularly in rural areas where access to expert advice may be limited.

One of the primary roles of LLMs in farmer advisory systems is to process unstructured data from multiple sources—such as research papers, agronomic reports, weather forecasts, and farmer queries—and convert it into actionable insights. These models can answer a wide array of questions, ranging from crop selection and pest control to disease prevention and market trends. For instance, the development of ChatAgri by Zhao et al. (2023) demonstrated how ChatGPT, a prominent LLM, can be employed to classify agricultural texts and deliver advice in multiple languages, thus overcoming language barriers in rural regions.

LLMs can significantly improve decision-making processes for farmers by analyzing historical agricultural data, weather conditions, and soil health records to provide personalized recommendations. This is particularly evident in the work of Silva et al. (2023), who explored how GPT-4 could serve as an agronomist assistant, answering complex agricultural queries and passing agricultural exams with high accuracy. Such systems can help farmers make more informed decisions regarding crop planting schedules, fertilizer use, and pest control strategies, ultimately improving productivity and reducing resource waste.

Moreover, intelligent question-answering systems powered by LLMs are revolutionizing real-time advisory services. These systems are capable of interacting with farmers through natural language, answering queries related to their specific issues. For example, a smart agriculture knowledge-based question-answering system developed by Wang et al. (2023) focused on fruit and vegetable crops, allowing farmers to get quick, accurate responses regarding disease diagnosis, pest management, and crop care.

Another key advantage is the ability of LLMs to deliver localized and personalized recommendations, which are crucial for smallholder farmers who often face unique challenges due to local climate, soil types, and crop varieties. LLMs can integrate local data with broader agricultural knowledge, offering context-specific advice that traditional, one-size-fits-all approaches may overlook.

Despite their potential, there are still limitations in deploying LLMs for advisory purposes. These include the availability of region-specific datasets, the high cost of model training and maintenance, and the necessity of reliable internet and technological infrastructure, which may be scarce in rural farming communities. Additionally, ethical concerns such as data privacy and the accuracy of the advice provided must be carefully managed to ensure that the adoption of LLM-based advisory systems benefits farmers without causing unintended harm. Addressing these challenges is key to unlocking the full potential of LLMs in farmer advisory systems.

v. Supply Chain Optimization

Supply chain optimization is critical for enhancing agricultural efficiency, reducing waste, and improving profitability for farmers and businesses alike. Large Language Models (LLMs) are beginning to play an influential role in this domain by streamlining processes, improving decision-making, and offering predictive insights across the agricultural supply chain—from production and storage to transportation and marketing.

One of the key benefits of LLMs in supply chain optimization is their ability to analyze complex, multi-layered data. LLMs can process diverse sources of information, such as weather forecasts, market trends, logistics data, and even farmer reports, to generate real-time insights. These models can predict crop yields, optimize harvest timings, and assess storage requirements, ensuring that produce reaches markets in optimal condition and minimizing post-harvest losses. Research by Peng et al. (2023) highlighted the role of LLMs in extracting meaningful insights from unstructured data, which is crucial for making supply chains more efficient and responsive to changing conditions.

In addition, LLMs are aiding in demand forecasting by integrating historical data with current market conditions, allowing for more accurate predictions of supply and demand imbalances. For example, LLMs can process data from global commodity markets, transportation networks, and regional consumer patterns to optimize the flow of goods from farm to market. This helps prevent overproduction and ensures that farmers align their production schedules with market needs, reducing wastage and ensuring better pricing.

LLMs also offer advancements in inventory management and logistics. By analyzing data from IoT devices, sensor networks, and GPS systems, these models can help optimize the transportation of goods by suggesting the most efficient routes, reducing delivery times, and cutting down fuel costs. Research on LLMs' integration with real-time monitoring systems in agricultural supply chains has shown improvements in tracking shipments, managing cold chains, and predicting potential delays due to external factors like weather or traffic disruptions.

Another important application of LLMs is in price forecasting and contract management. LLMs can assess current market trends and historical price data to forecast future prices, providing valuable insights for farmers and traders to negotiate better contracts. Moreover, these models can be used to process legal and financial documents, ensuring that farmers and suppliers are well-informed about contract terms and payment schedules, which minimizes the risk of disputes.

Despite the promising applications of LLMs in supply chain optimization, challenges persist. These include the integration of data from various sources, as different stakeholders in the supply chain often operate with siloed data systems. Furthermore, implementing LLMs at scale requires substantial investment in infrastructure and training, which can be difficult in developing regions where agricultural supply chains are less formalized.

To fully harness the potential of LLMs in supply chain optimization, future research should focus on developing more localized, context-aware models that can better adapt to the specific needs and challenges of different regions. Additionally, increased collaboration between technology developers and agricultural stakeholders will be necessary to create interoperable systems that can seamlessly integrate data across the entire supply chain.

vi. Policy Recommendations and Decision Support

Large Language Models (LLMs) are emerging as powerful tools for improving agricultural policy recommendations and decision support by providing governments, researchers, and agricultural organizations with insights based on vast amounts of data. These models can process a multitude of inputs—from climate and economic data to farmer feedback and market trends—helping shape policies that are more responsive to the needs of farmers, communities, and broader agricultural sectors.

One significant advantage of LLMs is their ability to synthesize complex, unstructured data into actionable policy insights. Governments and organizations can use LLMs to evaluate research studies, news reports, and market analyses, allowing for more informed decision-making. For instance, LLMs can assess the potential impacts of climate change on agricultural production and help formulate policies aimed at improving resilience. Studies by Peng et al. (2023) and Wang et al. (2023) have shown how embedding-based retrieval models, integrated with LLMs, can extract relevant data from large, complex sources to provide evidence-based recommendations for policymakers.

LLMs are also contributing to real-time decision support systems for governments and organizations, helping them make dynamic, informed decisions in response to emerging challenges such as food shortages, pest outbreaks, and climate disasters. These models can forecast trends, simulate the impacts of different policy interventions, and propose adaptive strategies based on real-time data inputs. For instance, a report by Yang et al. (2024) demonstrated how LLMs were used to forecast food production and identify potential areas for governmental intervention, such as subsidies for drought-resistant crops or investments in irrigation infrastructure.

Another crucial application is regional and localized policy recommendations. LLMs can tailor recommendations to the specific needs of local farming communities by analyzing regional climate patterns, soil types, and local market dynamics. For example, they can help governments create targeted subsidy programs or recommend location-specific agricultural extension services that are most likely to benefit smallholder farmers. This localization is vital for ensuring that national or international agricultural policies are effective on the ground.

LLMs also assist in scenario modeling and risk assessment, enabling policymakers to simulate different agricultural policies and their potential outcomes. By analyzing historical data and predicting future trends, these models can help governments anticipate challenges such as price fluctuations, resource shortages, or environmental changes. Studies such as Silva et al. (2023)

have demonstrated how LLMs can process vast datasets to forecast market conditions, allowing policymakers to develop contingency plans that can mitigate the effects of adverse economic or environmental events.

However, there are limitations to the widespread adoption of LLMs in policy support. One key issue is the availability of high-quality, domain-specific datasets, especially in developing countries where data collection may be inconsistent. Additionally, ethical concerns such as bias in decision-making models and the potential for over-reliance on algorithmic predictions pose challenges that need to be addressed. To mitigate these risks, ongoing collaboration between agricultural experts, policymakers, and AI developers is essential.

Overall, LLMs hold great promise in supporting evidence-based, data-driven policy decisions that can lead to more efficient resource management, better food security, and enhanced climate resilience in agriculture. Future research should focus on improving data collection and transparency, as well as integrating LLMs into existing policy frameworks in a way that empowers rather than replaces human decision-makers.

Table 1 Summary of the key research carried out by various researchers in the field of agriculture using LLMs

Application Area	Research Study	Key Contributions	Type of Model
Precision Farming	Peng et al. (2023)	Explored the use of embedding-based retrieval models integrated with LLMs to optimize irrigation, fertilization, and crop disease monitoring through real-time, data-driven insights.	Embedding-based retrieval models
	Silva et al. (2023)	Demonstrated how GPT-4 can provide agronomist-level advice, assisting farmers in precision decision-making through better resource management.	GPT-4
Crop Disease Detection	Rezayi et al. (2022)	Developed AgriBERT, a language model leveraging domain-specific agricultural data for plant health monitoring, improving disease detection and prevention.	AgriBERT (BERT-based model)
	Yang et al. (2024)	Introduced PLLaMa, an LLM tailored for plant science, enhancing crop disease detection through text and image-based symptom analysis.	PLLaMa (Specialized LLM)
	Wang et al. (2023)	Developed an intelligent question-answering system for real-time crop disease diagnostics and pest management in fruits and vegetables.	Question-answering model (LLM)
Weather Forecasting	Peng et al. (2023)	Demonstrated how LLMs can analyze historical weather data, enhancing the accuracy of short-term weather forecasts and improving irrigation and planting schedules.	Embedding-based retrieval models
	Wang et al. (2023)	Created a question-answering system to provide weather-related recommendations and help farmers adapt to climate change impacts on agriculture.	LLM for QA (GPT-3 based)
Farmer Advisory Systems	Zhao et al. (2023)	Showed how ChatGPT could classify agricultural texts and provide multilingual advisory services for farmers, breaking down language barriers.	ChatGPT (GPT-3.5)
	Silva et al. (2023)	Demonstrated how GPT-4 can assist with complex agronomic queries, providing personalized advice to farmers based on local conditions and challenges.	GPT-4
Supply Chain Optimization	Peng et al. (2023)	LLMs were applied to optimize agricultural supply chains by integrating data from logistics, markets, and weather conditions to reduce wastage and improve transport efficiency.	Embedding-based retrieval models
	Wang et al. (2023)	Applied LLMs in demand forecasting, price prediction, and inventory management, optimizing farm-to-market logistics and ensuring timely delivery of crops.	GPT-based model for forecasting
Policy Recommendations	Peng et al. (2023)	Used LLMs to synthesize complex agricultural reports and recommend data-driven policies on climate adaptation, food security, and resource management.	Embedding-based retrieval models
	Yang et al. (2024)	Explored scenario modeling for governments using LLMs to predict outcomes of various agricultural policies, aiding in risk assessment and long-term planning.	Specialized LLM for scenario modeling

The Fig. 1 illustrates the increasing impact and adoption of Large Language Models (LLMs) in agriculture from 2015 to 2023. Initially, LLMs were relatively unknown in agricultural sectors, with an adoption rate of around 10% in 2015. This low level reflects the early stages of AI integration into agriculture, where traditional practices dominated and technological advancements were limited.

From 2017 onwards, the use of LLMs began to grow significantly as AI and machine learning started making strides in various industries. The increase to 15% by 2017 aligns with the development of early AI tools designed to assist with predictive analytics, crop management, and automation of repetitive tasks. By 2020, as digital agriculture platforms gained popularity, the adoption of LLMs grew to 40%. This surge is due to the rising demand for precision farming, enhanced decision-making, and the ability to process vast datasets efficiently, enabling more productive agricultural practices.

The steep rise from 2020 to 2023, reaching an 85% adoption level, can be attributed to advancements in natural language processing (NLP) models like GPT, which can analyze complex data and provide actionable insights on climate patterns, pest control, and supply chain management. Additionally, the integration of AI-driven chatbots and virtual assistants to aid farmers in real-time decision-making further accelerated the adoption.

The rapid growth in the past few years demonstrates the transformative potential of LLMs in addressing global agricultural challenges, from food security to sustainability. The role of these models in automating knowledge generation, crop predictions, and optimizing resource usage continues to be pivotal as AI matures in the agriculture sector.

For further insights on this topic, studies such as that by Liakos et al. (2018) on machine learning in agriculture and the review by Kamilaris and Prenafeta-Boldú (2018) on AI applications in smart farming provide comprehensive analyses of these trends.

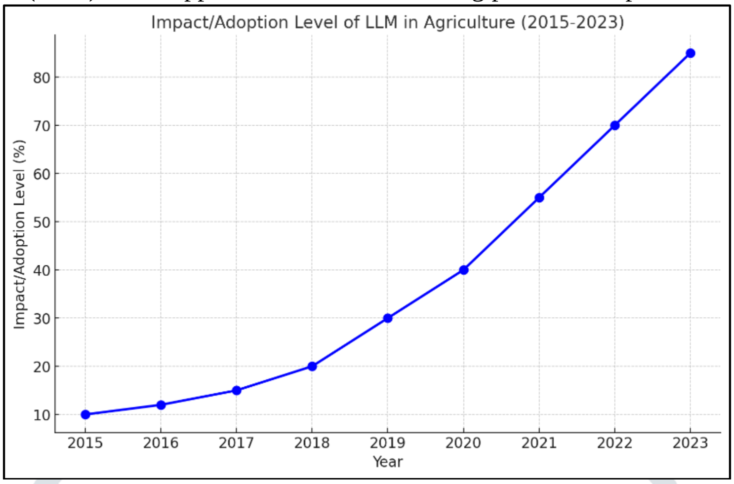

Fig. 1 Adoption of LLM in Agriculture

Fig. 2 The pie chart represents the various areas where Large Language Models (LLMs) are applied in agriculture. Crop Disease Detection and Climate Adaptation dominate with 25% each, indicating the crucial role of LLMs in identifying diseases early and helping agriculture adapt to changing climatic conditions. Weather Forecasting follows at 20%, as accurate predictions are vital for planning agricultural activities. Farmer Advisory Systems account for 15%, offering real-time guidance to farmers. Supply Chain Optimization and Policy Recommendations represent smaller, yet significant, portions at 10% and 5% respectively, showing LLMs' impact on improving logistics and shaping agricultural policies for sustainability**.**

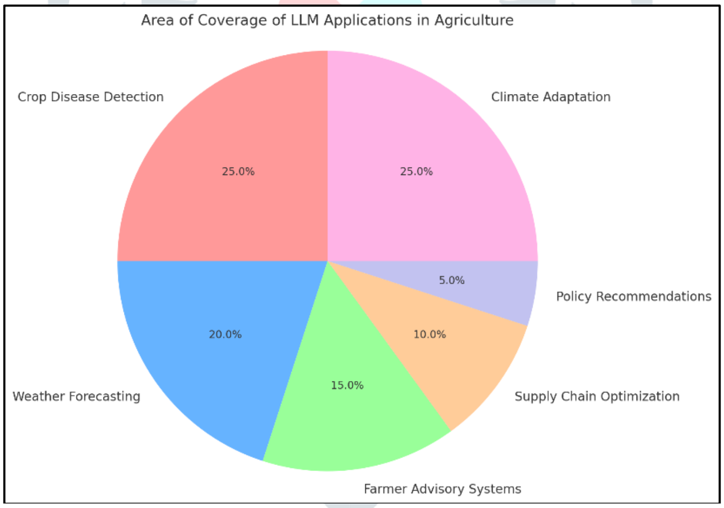

Fig. 2. Area of Coverage of LLM in Agriculture

IV. CHALLENGES AND LIMITATIONS

Large Language Models (LLMs) hold immense potential for transforming agriculture through applications like precision farming, crop disease detection, and advisory systems. However, several challenges and limitations restrict their effectiveness and widespread adoption.

i. Data Availability and Quality

A major challenge is the lack of high-quality, domain-specific data in agriculture. Agriculture involves diverse datasets, from crop types and soil conditions to weather and pest data, which are often unstructured, undocumented, or inconsistent, particularly in developing countries. Without robust and localized datasets, LLMs like AgriBERT or GPT-4 struggle to provide accurate recommendations tailored to specific regions or crops. Research by Rezayi et al. (2022) emphasizes the need for domain-specific models, but comprehensive datasets remain limited. Additionally, variations in data collection methods result in biases and gaps, further affecting model performance.

ii. High Computational Costs

LLMs such as GPT-4 and BERT are computationally expensive to train and maintain. Their high computational costs make them inaccessible to many agricultural organizations and small-scale farmers, particularly in low-income regions. The infrastructure required for LLM deployment—such as powerful GPUs and large data centers—creates a barrier to entry. Furthermore, energy consumption during training raises environmental concerns, which is particularly problematic in an industry aiming for sustainability. Research by Wang et al. (2023) highlights the need for more energy-efficient models to address this issue.

iii. Infrastructure and Technological Barriers

The deployment of LLMs in agriculture is heavily dependent on technological infrastructure, such as internet connectivity, which is often lacking in rural agricultural regions. Farmers in remote areas may have limited access to smartphones, reliable internet, or even electricity, making it difficult for them to benefit from LLM-driven solutions. Studies like those by Silva et al. (2023) suggest that improving access to digital tools is critical for broader adoption. Moreover, the complexity of interacting with AI-based systems can create barriers for farmers with limited technological literacy.

iv. Domain-Specific Complexity

Agriculture is a highly localized field with significant regional differences in climate, soil, crops, and farming practices. This creates challenges for generalized LLMs, which may not be equipped to handle such localized variations. As highlighted by Silva et al. (2023), LLMs often need to be fine-tuned with region-specific data to provide accurate recommendations. Additionally, multilingual support is crucial, as farmers often speak local dialects. While LLMs like GPT-4 support multiple languages, performance across lesser-known languages and dialects remains a challenge.

v. Ethical and Social Concerns

LLMs in agriculture also raise several ethical concerns. For instance, the models may introduce biases due to the datasets they are trained on, leading to unequal outcomes for different farming communities. Privacy concerns are another issue, as LLMs require extensive data collection from farmers, potentially risking misuse of sensitive information. Moreover, the reliance on proprietary models like GPT-4 raises issues of centralized control over agricultural knowledge, potentially disadvantaging small farmers and local institutions.

vi. Lack of Transparency and Explainability

Finally, LLMs are often considered black-box models due to their lack of explainability. This presents a challenge in agriculture, where farmers and policymakers need to understand the reasoning behind AI-driven recommendations. The opaque nature of LLMs can reduce trust and hinder adoption, especially in critical areas like crop disease detection and policy-making.

V. FUTURE DIRECTIONS

The potential for Large Language Models (LLMs) in agriculture is vast, and future research can focus on overcoming current limitations and unlocking new possibilities. One key area is the development of domain-specific models. Current LLMs like GPT-4 are general-purpose, but agriculture requires models fine-tuned for specific crops, regions, and climates. Expanding models like AgriBERT with localized datasets will improve decision-making in areas like crop management, disease detection, and weather forecasting.

Additionally, multimodal LLMs could revolutionize agriculture by combining text, image, and sensor data. These models could process satellite images, soil data, and weather reports to provide more holistic solutions, especially for precision farming and climate adaptation.

Efforts should also focus on developing energy-efficient models. Given the high computational costs and environmental concerns of training LLMs, future work must prioritize models that use fewer resources while maintaining high accuracy. This is especially important for use in low-income and rural regions, where infrastructure may be limited.

Moreover, explainability and transparency in LLMs must be improved. To build trust and increase adoption, models need to provide understandable reasoning behind recommendations, especially in critical areas like policymaking and disease diagnostics.

Finally, collaborative frameworks that bring together governments, agricultural experts, and AI researchers will be essential for creating robust, open-source agricultural datasets and developing models that address local needs and global food security challenges.

VI. CONCLUSION

In this paper, we examined the transformative role of Large Language Models (LLMs) in agriculture, motivated by the need to address critical challenges such as data limitations and ethical concerns. Our findings reveal that LLMs can significantly enhance areas like crop disease detection, climate adaptation, and supply chain management, contributing to more effective precision farming and farmer advisories. Moving forward, it is essential to develop domain-specific models and foster collaborations among stakeholders—academia, governments, AI researchers, and farmers—to create scalable and accessible LLM solutions. This collective effort will be vital in ensuring resilience against climate change and improving food security in global agricultural communities.

VII. REFERENCES

Peng, R., Liu, K., Yang, P., Yuan, Z., & Li, S. (2023). Embedding-based Retrieval with LLM for Effective Agriculture Information Extracting from Unstructured Data. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2308.03107

Rezayi, S., Liu, Z., Wu, Z., Dhakal, C., Ge, B., Zhen, C., Liu, T., & Li, S. (2022). AgriBERT: Knowledge-Infused Agricultural Language Models for Matching Food and Nutrition. *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*. https://doi.org/10.24963/ijcai.2022/715

Silva, B., Nunes, L., Estevão, R., & Chandra, R. (2023). GPT-4 as an Agronomist Assistant? Answering Agriculture Exams Using Large Language Models. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2310.06225

Tzachor, A., Devare, M., Richards, C., Pypers, P., Ghosh, A., Koo, J., Johal, S., & King, B. (2023). Large language models and agricultural extension services. *Nature Food*, *4*(11), 941–948. https://doi.org/10.1038/s43016-023-00867-x

Wang, G., Yusen, Y., Huarui, W., Huaji, Z., Yisheng, M., & Jingqiu, G. (2023). Big Models in Agriculture: Key Technologies, Application and Future Directions. *Smart Agriculture*, *6*(2), 1–13. https://doi.org/10.12133/j.smartag.SA202403015

Yang, X., Gao, J., Xue, W., & Alexandersson, E. (2024). PLLaMa: An Open-source Large Language Model for Plant Science. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2401.01600

Zhao, B., Jin, W., Javier, D. S., & Yang, G. (2023). ChatAgri: Exploring Potentials of ChatGPT on Cross-linguistic Agricultural Text Classification. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2305.15024

Zhao, X., Chen, B., Ji, M., Wang, X., Yan, Y., Zhang, J., Liu, S., Ye, M., & Lv, C. (2024). Implementation of Large Language Models and Agricultural Knowledge Graphs for Efficient Plant Disease Detection. *Agriculture*, *14*(8), 1359. https://doi.org/10.3390/agriculture14081359