Assessing the Usability, Diagnostic Performance and Adoption Potential of Crop Cura AI-Based Diagnosis Tool Among Agriculture Students in Nnamdi Azikiwe University, Nigeria

Victor Onyekachukwu[1], Soetan Olatunbosun Jonathan[2], Christian Iwuagwu[3], and Akerele Dare

Crop Cura Limited[1], Department of Agricultural and Rural Development, Federal University of Agriculture, Abeokuta[2], Department of Crop Sciences, Nnamdi Azikiwe University[3], Department of Crop Sciences and Horticulture[4]

Abstract

This study assessed the usability, perceived diagnostic performance, trust, practical usefulness and adoption potential of an artificial intelligence (AI)-based crop-diagnosis application among agriculture-related students in Nigeria. The analysis used 222 questionnaire and combined frequency distributions and percentages with an ordinary least squares multiple linear probability model using HC3 heteroskedasticity-robust standard errors. Most respondents had not previously used a farming application (81.5%), although 54.1% reported farming outside school. Initial usability was generally favourable: 85.1% rated setup as easy or very easy, and 81.5% rated navigation as easy or very easy. However, 47.3% experienced an application failure, freezing or photo-upload problem, while 68.5% reported that network connectivity or data costs affected use at least sometimes. Perceived diagnostic agreement was moderate to positive: 58.1% said the diagnosis matched their visual observations most times or always, while 9.9% reported that it rarely or never matched. Most respondents reported mostly or complete trust (76.1%), had applied the recommendation as provided or with an organic alternative (85.6%), and perceived clear or slight crop improvement (86.5%). Overall, 80.6% would recommend the application to other farmers. In the multivariable model, perceived diagnostic agreement ($B = 0.094$, $p = .006$) and trust ($B = 0.170$, $p = .001$) were positively associated with willingness to recommend; ease of navigation, technical failure, network/data burden and perceived crop improvement were not statistically significant at the 5% level. The findings indicate promising user acceptance but also substantial reliability and access concerns. Responsible deployment should therefore prioritise field validation on locally important crops, transparent communication of uncertainty, affordable and safe recommendations, low-data/offline functionality, and integration with agricultural extension services.

Introduction

The specific research gap addressed here is the need to examine not only whether users find an AI crop-diagnosis tool easy to operate, but also how perceived diagnostic agreement and trust relate to endorsement while technical access barriers are considered. The study therefore provides exploratory evidence on user experience and recommendation intention, while explicitly separating perceptions from independently verified diagnostic performance.

Artificial intelligence (AI), particularly computer vision and deep learning, is increasingly being investigated as a way to identify visible crop symptoms from photographs and provide rapid preliminary guidance. Such tools could help users recognise possible plant-health problems earlier, especially where access to timely diagnostic expertise is limited. Nevertheless, the practical value of an AI crop-diagnosis

application depends on more than its performance on a benchmark dataset. The quality and diversity of images, crop variety, symptom stage, lighting, background clutter, camera quality, and the distinction between disease and non-infectious stress can all affect the reliability of an output. Recent reviews emphasise that models trained or tested under controlled conditions may not generalise well to heterogeneous field environments, and that datasets can be geographically or biologically unrepresentative (Jafar et al., 2024; Khan et al., 2025; Shafay et al., 2025). These limitations make independent, local field evaluation essential before diagnostic claims are used to guide consequential farm decisions.

Usability and adoption are also shaped by infrastructure and the fit between a tool and the everyday realities of intended users. In this survey, respondents were agriculture-related students rather than a representative sample of Nigerian smallholder farmers, but their experiences provide useful early evidence about interface accessibility, technical reliability and perceived usefulness. Nigerian research on generative AI among small-scale farmers reports growing interest alongside barriers involving awareness and access to digital devices, while conventional information channels remain important (Shitu et al., 2025). This context suggests that an application that requires repeated high-resolution uploads, stable internet access or substantial digital confidence may exclude some intended users even if its interface appears simple during a classroom exercise. Low-bandwidth operation, data-efficient image handling, local-language and voice support, and extension-worker assistance should therefore be treated as core design considerations rather than optional additions.

Trust is particularly important because users may act on recommendations that involve spending money, applying pesticides or changing crop-management practices. A convincing interface does not guarantee a correct diagnosis, and an incorrect recommendation can lead to unnecessary expenditure, crop damage, unsafe chemical use or delayed expert intervention. Recent work highlights the importance of field-ready evaluation, interpretability, local datasets and deployment design that fits resource-constrained environments (Khan et al., 2025; Shafay et al., 2025). AI crop diagnosis is consequently best positioned as a decision-support aid that helps users formulate questions and identify possible next steps, not as an unquestionable replacement for plant-health specialists or agricultural extension officers.

Against this background, the present study examines user experience with an AI-based crop-diagnosis application. It describes prior exposure to farming apps, farming background, ease of setup and navigation, technical failures, connectivity/data burden, perceived agreement between AI outputs and visual observations, trust, use of recommendations, perceived crop response, treatment affordability and willingness to recommend the tool. It also estimates a multivariable model of recommendation intention to identify which user perceptions are statistically associated with that outcome when other included variables are held constant.

.

Data and analytical methods

Descriptive analysis used frequency counts and percentages. Categories were cleaned by trimming leading/trailing spaces where relevant. Open-ended answers were considered qualitatively in the

interpretation, but were not subjected to formal thematic coding. For regression, the outcome was willingness to recommend the tool (Yes = 1; “Not yet” and “No” = 0). Predictors were coded as ordered scores: diagnosis agreement (Never=1 to Always=5), trust (Not at all=1 to Completely=4), ease of navigation (Very confusing=1 to Very easy=4), network/data burden (No=0, Sometimes=1, Yes often=2), crop improvement (worse=1 to clearly improved=5), and technical failure (Yes=1, No=0). An ordinary least squares (OLS) multiple linear probability model was fitted with HC3 heteroskedasticity-robust standard errors. This is a pragmatic, interpretable model for a binary outcome; predicted probabilities may fall outside 0–1, and a logistic regression would be a useful robustness check. Ordinal coding assumes equal spacing between response categories, so coefficients should be interpreted cautiously.

Table 1. Previous experience

Most respondents (81.5%) had not previously used a farming application, indicating that the survey exercise introduced many participants to this type of digital agricultural service.

Table 1: Previous experience

Response category	Frequency (n)	Percentage (%)
No	181	81.5
Yes	41	18.5

App usability and technical access

Prior experience and usability

Setup was rated easy or very easy by 85.1%, and navigation was rated easy or very easy by 81.5%. These results suggest that the basic interaction design was accessible to many first-time users. This is encouraging for technology transfer because a complicated onboarding process can discourage use before users have a chance to evaluate the quality of the advice.

However, positive usability ratings should not be interpreted as proof that the application is diagnostically dependable. Usability describes how readily users can operate a system; diagnostic validity concerns whether the system correctly identifies a disease or other plant-health condition. A tool can be simple to navigate yet still generate uncertain or inaccurate outputs. Reviews of AI plant-disease detection identify limited dataset diversity, class imbalance and the gap between controlled image collections and real field photographs as continuing concerns (Jafar et al., 2024; Khan et al., 2025). The practical implication is that interface refinement and diagnostic validation should proceed together. Future evaluation should include participants with different levels of digital literacy and farming experience, record task completion and error rates, and test whether users understand the output and its uncertainty. The 81.5% of respondents who had never used a farming app also indicates a need for short demonstrations, visual instructions and extension-led orientation before broader deployment.

Table 2: App usability and technical access

Response category	Frequency (n)	Percentage (%)
Easy	133	59.9
Very easy	56	25.2
I needed help	24	10.8
Difficult	9	4.1
Ease of navigation		
Easy	128	57.7
Very easy	53	23.9
Confusing in places	39	17.6
Very confusing	22	0.9

Reliability and connectivity

A major implementation concern is the frequency of technical problems: 105 respondents (47.3%) reported that the application failed, froze or encountered photo-upload problems. In addition, 68.5% reported that network access or data costs affected their use at least sometimes, including 18.0% who said this happened often. These are not minor inconveniences. A failed upload can interrupt diagnosis at the point when a user needs advice, while repeated attempts may consume data, reduce confidence and cause users to abandon the tool. In rural and peri-urban settings, reliable operation must be considered part of the service's practical performance.

The findings are consistent with recent reviews identifying deployment constraints, environmental variation and the need for offline or low-resource operation as important challenges for crop-disease AI (Khan et al., 2025; Shafay et al., 2025). Product development should therefore include image compression, upload-resume functions, clear error messages, compatibility with lower-cost smartphones and testing under variable network conditions. Where technically feasible, lightweight on-device diagnosis or offline storage with later synchronisation could reduce dependence on continuous connectivity. The application should also distinguish a technical failure from a diagnostic result: users need to know whether an image was processed successfully and should not be left to interpret an incomplete response as a diagnosis. Future studies should measure the frequency and causes of failure, upload time, data consumed per diagnosis and success rates across different devices and network conditions. The present data show that technical problems were common, but the regression did not find a statistically significant independent association between reported failure and recommendation intention. This does not make reliability unimportant; it means that, in this sample and model, failure was not a statistically significant predictor after accounting for the other included perceptions.

Table 3: Reliability and connectivity

Response category	Frequency (n)	Percentage (%)
No	117	52.7
Yes	105	47.3

Network/data burden		
Sometimes	112	50.5
No	70	31.5
Yes, often	40	18.0

Perceived diagnosis agreement and crop differences

A total of 129 respondents (58.1%) reported that the diagnosis matched their own visual observations most times or always. Conversely, 22 respondents (9.9%) selected rarely or never, while 32.0% selected sometimes. This distribution suggests that users often perceived the output as plausible, but consistency was not universal. Respondents also reported perceived differences between crops: 23.0% considered the application better on okra and 15.8% better on ugu, while 31.1% perceived performance as about the same and 30.2% monitored only one crop. Open-ended comments mentioning crop misidentification or unknown outputs reinforce the need to assess performance by crop and condition rather than assuming that an application works equally well across all plant species.

These figures are measures of perceived agreement, not objective diagnostic accuracy. Respondents' visual observations were not independently confirmed by plant pathologists, laboratory tests or another reference standard. Agreement between a user and an AI system can occur even when both are mistaken, and some plant symptoms can be caused by disease, nutrient deficiency, drought, pests or physical injury. Recent reviews stress that changes in lighting, background, cultivar, disease stage and field context can reduce the transferability of models trained on curated datasets (Jafar et al., 2024; Khan et al., 2025; Shafay et al., 2025). A robust validation study should collect field images across multiple locations and seasons, obtain expert-confirmed labels, include healthy plants and look-alike symptoms, and report crop-specific sensitivity, specificity, precision, recall and confusion matrices. Performance should also be assessed across disease severity levels and common Nigerian crops, including okra and ugu, rather than relying on one overall satisfaction score. Where the model is uncertain or a symptom is outside its supported classes, the safest response is to state the limitation and direct the user to an extension officer or plant-health expert.

Table 4: Perceived diagnosis agreement and crop differences

Response category	Frequency (n)	Percentage (%)
Most times	102	45.9
Sometimes	71	32.0
Always	27	12.2
Rarely	15	6.8
Never	7	3.2
Perceived crop-specific difference		
About the same	69	31.1
I only monitored one	67	30.2
Better on Okra	51	23.0
Better on Ugu	35	15.8

Trust, recommendation

Trust was relatively high: 76.1% of respondents reported mostly or complete trust in the diagnosis, while 23.0% trusted it a little and 0.9% not at all. In total, 179 respondents (80.6%) said they would recommend the application to farmers.

Table 5: Trust and use of recommendations

Response category	Frequency (n)	Percentage (%)
Mostly	130	58.6
A little	51	23.0
Completely	39	17.6
Not at all	2	0.9
Total	222	100.0

Application, crop response and affordability

A large majority of respondents (85.6%) said they had applied the recommendation either exactly as provided (61.7%) or by using an organic alternative (23.9%). In addition, 86.5% reported clear or slight crop improvement, while 7.2% said it was too early to tell or that there was no change, and 2.7% reported that the crop got worse. These responses suggest that participants generally perceived the recommendations as actionable and often associated their use with a positive crop response. However, the survey did not include untreated comparison plots, baseline severity measurements, objective yield data or independent assessment of plant recovery. It is therefore not possible to conclude that the application caused the reported improvement. Weather, irrigation, fertiliser, natural disease progression, other treatments and differences in observation periods may all have influenced crop condition.

Treatment feasibility was rated positively by 84.2% of respondents, who considered the recommended solution affordable or obtainable. Nevertheless, 15.3% were unsure or found it difficult, indicating that affordability and local availability were not universal. A technically correct recommendation may still have limited value if the product is unavailable locally, too expensive, unsuitable for the crop or unsafe when applied incorrectly. Advice shosuld therefore be tailored to locally available inputs and should explain dosage, protective measures, waiting periods and when professional advice is required. Where appropriate, recommendations should include integrated pest management and non-chemical options rather than defaulting to chemical treatment. The Nigerian study by Shitu et al. (2025) similarly underscores that digital readiness and access barriers need attention when considering AI-supported agricultural information services.

Future evaluation should distinguish between user satisfaction, treatment adherence and verified agronomic effectiveness. Researchers could compare AI-assisted advice with standard extension advice and usual farmer practice using replicated field trials, expert diagnosis and objective indicators such as symptom progression, marketable yield, input expenditure and treatment costs. Such designs would establish whether the perceived usefulness observed here translates into measurable agronomic and economic benefits. Until then, the present findings are best interpreted as encouraging user-reported experiences, not evidence of a causal yield or disease-control effect.

Table 6: Application, crop response and affordability

Response category	Frequency (n)	Percentage (%)
Yes, exactly as given	137	61.7
Yes, but I used an organic alternative	53	23.9
No	32	14.4
Perceived crop response		
Yes clearly	123	55.4
A little	69	31.1
Too early to say	14	6.3
No change	10	4.5
It got worse	6	2.7
Affordability/availability of recommendation		
Yes	187	84.2
Not sure	27	12.2
Difficult	7	3.2
No	1	0.5

Readiness for smallholder-farmer use

Open-ended responses were varied, but a recurring position was that the application could be useful as a support tool while not replacing extension officers or plant-health specialists. Respondents' suggested improvements included better crop and disease identification, offline functionality, local-language or voice features, clearer instructions and treatment options that reflect local cost and availability. These recommendations align with recent research calling for field-representative datasets, more generalisable models and practical deployment strategies, including multilingual and offline support (Khan et al., 2025; Shafay et al., 2025). The comments are especially relevant because adoption is not simply a matter of whether a user likes the application; sustained use also depends on whether it works when needed, gives understandable advice and fits local farming practices.

Readiness for smallholder use should consequently be understood as conditional rather than established. The survey involved students and course participants, and their responses cannot be treated as a representative measure of willingness among Nigerian smallholder farmers. Farmers may face different constraints in device ownership, literacy, language, internet access, ability to pay for recommended inputs and confidence in digital services. Shitu et al. (2025) found that Nigerian small-scale farmers' engagement with generative AI coexists with uneven awareness and access to devices, reinforcing the need to design for inclusion. Co-design with farmers, extension agents, crop specialists and local-language speakers would help ensure that the tool's workflow and advice are relevant to real production contexts.

A practical deployment model would combine AI-assisted preliminary screening with human support. The app could identify likely conditions, ask for additional photographs when image quality is poor, display confidence or uncertainty in plain language, and provide a route to contact an extension officer when a

result is uncertain or the recommended intervention has substantial cost or safety implications. Extension agents could also use aggregated, privacy-protective data to identify recurring crop-health concerns, provided data governance and farmer consent are addressed. Before large-scale rollout, pilot programmes should test diagnostic performance, accessibility, user comprehension, repeat use, affordability and safety across several locations and farming seasons. This staged approach would enable promising features to be improved while reducing the risk that users interpret an unverified AI output as a definitive diagnosis.

Table 7: Adoption potential

Response category	Frequency (n)	Percentage (%)
Yes	179	80.6
Not yet	41	18.5
No	2	0.9

Hypothesis

The regressions results help clarify this pattern. Holding the other included predictors constant, a one-category increase in perceived diagnostic agreement was associated with an estimated 9.4-percentage-point increase in recommendation probability (B = 0.094, p = .006; 95% CI [0.027, 0.160]). A one-category increase in trust was associated with an estimated 17.0-percentage-point increase (B = 0.170, p = .001; 95% CI [0.074, 0.266]). Both confidence intervals exclude zero, supporting positive statistical associations in this sample. The model was statistically significant overall (robust Wald/F = 7.721, p < .001) and explained approximately 20.0% of the variation in recommendation (R^2 = .200; adjusted R^2 = .178).

By contrast, ease of navigation, technical failure, network/data burden and perceived crop improvement were not statistically significant at the 5% level in this model. This should not be read as evidence that these factors have no practical relevance. Statistical significance depends on sample composition, measurement quality, predictor overlap and model specification; a non-significant coefficient may still be operationally important. The positive associations for trust and perceived agreement are plausible because users may be more willing to endorse a tool when its outputs seem consistent with what they observe and when they consider the advice credible. However, the cross-sectional survey cannot establish that trust or perceived agreement causes recommendation. Responses may be influenced by prior expectations, the novelty of the tool, social desirability or shared measurement methods.

The outcome was binary, while several predictors were ordinal scores treated as if adjacent response categories were equally spaced. The OLS linear probability model offers an accessible interpretation in percentage-point terms, but it can produce predictions outside the 0–1 range and its estimates depend on the coding assumptions. A logistic regression, alternative ordinal coding and sensitivity analysis would strengthen a journal submission. Future work should also assess whether the associations differ by department, prior farming experience, crop tested or previous app use. Overall, the strongest practical message is that trust should be earned through demonstrated and transparent reliability, not promoted

through persuasive language alone. Confidence indicators, explanations of possible causes, explicit uncertainty messages and clear referral pathways may help users calibrate trust appropriately

Table 8: Regression results

Predictor	B	Robust SE	p-value	95% CI lower	95% CI upper
Constant	0.088	0.213	0.680	-0.329	0.504
Diagnosis agreement (1–5)	0.094	0.034	0.006	0.027	0.160
Trust (1–4)	0.170	0.049	0.001	0.074	0.266
Ease of navigation (1–4)	0.014	0.038	0.719	-0.061	0.088
Technical failure (yes=1)	-0.053	0.053	0.313	-0.157	0.050
Network/data burden (0–2)	-0.009	0.038	0.818	-0.084	0.067
Perceived crop improvement (1–5)	-0.028	0.025	0.251	-0.077	0.020

Model fit: R^2=0.200; adjusted R^2=0.178; robust Wald/F statistic=7.721; model p-value=1.579e-07; N=222.

Conclusion

The study suggests that users generally found the application navigable and many were willing to recommend it, but technical failures, network/data constraints and inconsistent crop diagnosis remain important concerns. The regression indicates that perceived diagnostic agreement and trust are the clearest statistical correlates of willingness to recommend the tool in this sample. The most defensible conclusion is that AI crop diagnosis has promising decision-support potential, but deployment for Nigerian smallholder farmers should be accompanied by local field validation, offline/low-data design, affordable recommendations and integration with agricultural extension services.

Recommendations

Validate the app against expert-confirmed diagnoses using field photographs of locally important crops, including okra and ugu, across multiple locations and growing conditions. Report crop-specific sensitivity, specificity, precision, recall and confusion matrices rather than relying only on user satisfaction.

References

Jafar, A. J., Bibi, N. N., Naqvi, R. A., Sadeghi-Niaraki, A., & Jeong, D. (2024). Revolutionizing agriculture with artificial intelligence: Plant disease detection methods, applications, and their limistations. *Frontiers in Plant Science, 15,* Article 1356260. https://doi.org/10.3389/fpls.2024.1356260

Khan, S. U., Alsuhaibani, A., Alabduljabbar, A., Almarshad, F., Altherwy, Y. N., et al. (2025). A review on automated plant disease detection: Motivation, limitations, challenges, and recent advancements for future research. *Journal of King Saud University Computer and Information Sciences, 37,* Article 34. https://doi.org/10.1007/s44443-025-00040-3

Shitu, A., Anafi, S., Tulagha, I., Nain, M. S., Ojobola, F., Olaniyan, O., Alabi, O., Ayegbusi, O., Bamigboye, O., Olatunji, C., Fanu, A., Ayotunde, K., Makinde, O., Shitu, M., Gabriel, G., Dandara, B., & Adeyowin, O. B. (2025). Use of generative AI by small-scale farmers in Nigeria: An empirical study. *Indian Journal of Extension Education, 61(4*). https://doi.org/10.48165/IJEE.2025.61424

Upadhyay, A., Chandel, N. S., Singh, K. P., & Chakraborty, S. K. (2025). Deep learning and computer vision in plant disease detection: A comprehensive review of techniques, models, and trends in precision agriculture. *Artificial Intelligence Review, 58,* Article 92. https://doi.org/10.1007/s10462-024-11100-x

Shafay, M., Hassan, T., Owais, M., Hussain, I., Khawaja, S. G., Seneviratne, L., & Werghi, N. (2025). Recent advances in plant disease detection: Challenges and opportunities. Plant Methods, 21, Article 140. https://doi.org/10.1186/s13007-025-01450-0