Computers and Education Open 10 (2026) 100369

Contents lists available at ScienceDirect

Computers and Education Open

journal homepage: www.sciencedirect.com/journal/computers-and-education-open

Developing future workforce skills through remote practical training☆

Judith Nkechinyere Njoku [a], Patience Chizoba Mba [a], Senorpe Asem-Hiablie [b], Daniel Dooyum Uyeh [a,*]

[a] Department of Biosystems and Agricultural Engineering, Michigan State University, East Lansing, Michigan, USA
[b] Biotechnology Department, Shell International Exploration and Production Inc.Shell Technology Center, Houston, USA

ARTICLE INFO

Keywords:
Internship
Remote learning
Workforce
Training

ABSTRACT

The coronavirus disease 2019 (COVID-19) pandemic accelerated the global shift to remote work, highlighting a critical need for robust digital training environments. This paper presents a remote internship program in artificial intelligence (AI) and machine learning (ML) designed as a computer-supported experiential learning model to equip learners, primarily based in Africa, with essential digital skills. Spanning three iteratively designed cohorts, the program utilized a suite of digital platforms to facilitate tool-mediated mentorship, collaborative coding, and asynchronous feedback. Employing a mixed-methods evaluation, this study investigates how digital learning environments and a dual mentorship model support advanced technical skill acquisition. The program demonstrated high completion rates and significant post-internship outcomes, with over half of the participants advancing to graduate school or securing technical roles. This paper discusses the program's technology-mediated structure, cohort evolution, and lessons learned, offering reproducible insights into designing digital learning infrastructures for remote workforce development.

1. Introduction

1.1. Background

Due to the COVID-19 pandemic, a global trend toward increasing remote work emerged [1]. This shift was vital for the survival of numerous industries during the pandemic and transformed the working environment by popularizing two work methods, remote and hybrid, in addition to on-site work. The hybrid approach permitted employees to divide work hours between remote and on-site activities, depending on the circumstances. The remote workforce in various industries underwent a drastic percentage change following the lockdown. For instance, according to a survey by *Great Place to Work*, there was a 77% increase in the remote workforce of industries offering professional services [2]. Similarly, the information technology industry experienced a 64% increase in the remote workforce after the pandemic [2]. Notably, the healthcare industry experienced the least significant increase, at 12%, given the requirements and risks associated with having healthcare professionals work remotely [2]. The detailed result of this survey is illustrated in Fig. 1. A survey by Gartner conducted in 2020 revealed that 88% of organizations mandated their employees to work from home as the virus spread globally [3]. Another study by the U.S. Bureau of Labor Statistics revealed that 80% of employees in specific sectors continued to work remotely [4]. Currently, numerous companies operate on a fully remote or hybrid basis. With a substantial portion of the workforce working remotely, concerns have rose regarding the level of preparedness of learners for remote work. Additionally, UNESCO reported a global increase in school closures due to inactivity during the pandemic, affecting approximately 1.38 billion learners, including those at the tertiary education level [5].

Unfortunately, despite the successful implementation of remote working arrangements in the workforce, not all educators adopted similar approaches to mitigate these drastic closures [6]. As a result, many students were left idle, becoming delinquent or stagnant in their learning [7]. There was a significant increase in attendance on massive open online course (MOOC) platforms such as Coursera and Edx during the pandemic, indicating the need for more remote learning opportunities [8].

A study by *McKinsey* analyzing 800 occupations, 2000 tasks, and spanning nine countries [9] revealed the following:

- The feasibility of remote work is dictated by the nature of tasks and activities rather than by the specific occupations

☆ This article is part of a Special issue entitled: 'AI in African Edu' published in Computers and Education Open.

* Corresponding author.
E-mail addresses: judithnjoku24@gmail.com (J.N. Njoku), patchizzymba@gmail.com (P.C. Mba), senorpe.hiablie@shell.com (S. Asem-Hiablie), uyehdani@msu.edu (D.D. Uyeh).

https://doi.org/10.1016/j.caeo.2026.100369
Received 10 October 2025; Received in revised form 15 March 2026; Accepted 8 May 2026
Available online 20 May 2026

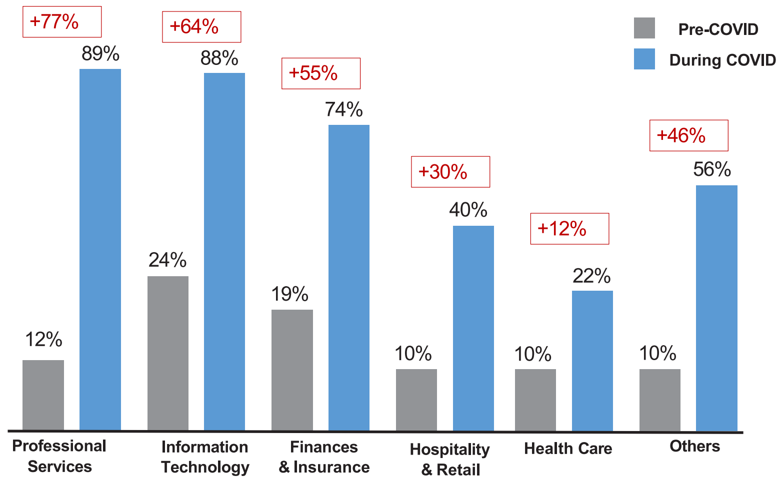

Fig. 1. Estimated workforce distribution by industry before and during COVID-19, illustrating shifts in employment patterns and percentage growth across occupations [2]. The figure contextualizes the increasing demand for digitally adaptable roles and remote-capable skill sets that motivate the development of virtual AI/ML training programs.

- The activities that had the highest potential for remote work included interaction with computers and updating knowledge.
- The sectors of finance, management, professional services, and information technology exhibited the highest potential for remote work.
- Workforces in advanced economies had greater opportunities to engage in remote work compared to those in emerging economies.
- While the majority of the workforce could not work remotely, up to one quarter in advanced economies can do so three to five days a week.

Based on this study in [9], and illustrated by Fig. 2, highlighting the percentage of workforce number of days per week of potential remote work without productivity loss, a significant number of occupations can effectively thrive by working remotely. Many industries embraced the idea of remote work post-COVID and have been recruiting accordingly [10]. Given this growing trend of remote work in various industries, preparing learners worldwide for such scenarios is crucial to ensuring their productivity despite remote working arrangements.

With the significant shift in work modalities, there is a profound implication for internship programs, traditionally designed around in-person experiences that facilitate direct mentorship and hands-on learning. The need to adapt remote and hybrid work models has spurred innovative approaches to internship design, ensuring that interns continue to gain invaluable industry experience, even in fields where remote work was previously unfathomable. For instance, remote internships have become the norm in the professional services and IT sectors, where the adoption of remote work has been exceptionally high [11]. Moreover, the cost efficiencies associated with remote internships, such as the reduced need for physical office space and commuting, can enable organizations to allocate more resources toward enhancing the internship experience, perhaps by offering stipends, investing in better training resources, and designing a well-structured internship program that maximizes learning and productivity and exposes them to cutting-edge learning technologies [12].

1.2. Remote internships as computer-supported experiential learning

Despite the rapid adoption of remote work, there remains a critical technology learning gap regarding how computing systems can effectively structure experiential learning [13,14]. Much of the existing discourse treats remote internships merely as logistical alternatives to in-person roles [7]. However, as will be explored in the subsequent literature review, there is a pressing need to treat remote internships as a learning technology problem [6]. By framing the internship as a computer-supported collaborative learning (CSCL) environment, digital tools are positioned not just as communication utilities, but as vital infrastructure that actively mediates mentorship, accountability, and knowledge construction [15,16].

Furthermore, unlike other STEM disciplines that require physical laboratories, AI workforce development is uniquely suited for this remote experiential model, as computational and knowledge-updating tasks consistently exhibit the highest potential for remote adaptation [9]. During the internship, learners based primarily in Africa were provided with high-compute resources and multimodal agricultural datasets (such as soil images and IoT sensor data) that were not generated in their local regions. This demonstrates the profound viability of remote AI work: as long as learners possess internet connectivity and cloud-based computational access, geographical boundaries to cutting-edge research and workforce training are effectively eliminated.

1.3. Objectives

This study aims to provide insights into an annual remote internship program (six-month duration), which has spanned 3 cohorts. The main objectives of this paper include the following:

- To present the remote internship program, detailing the innovative methods employed.
- To showcase the methods, implementation strategies, tools, and technologies that facilitated the internships.
- To demonstrate the program's impact on expanding access to professional development opportunities, fostering inclusivity, and building a resilient future workforce equipped with the skills for the digital age,
- To contribute to the broader discourse on the best practices for remote internships and the continuous improvement of the remote internship scheme worldwide.

To better address these objectives, the following questions will be answered by discussing details of our internship being focused on artificial intelligence (AI) and machine learning (ML):

1. What specific AI/ML skills were best acquired through remote training, and how does this method compare to traditional in-person learning?
2. What challenges and opportunities might arise when nurturing AI/ML talents through remote platforms, and how could these challenges be effectively addressed?

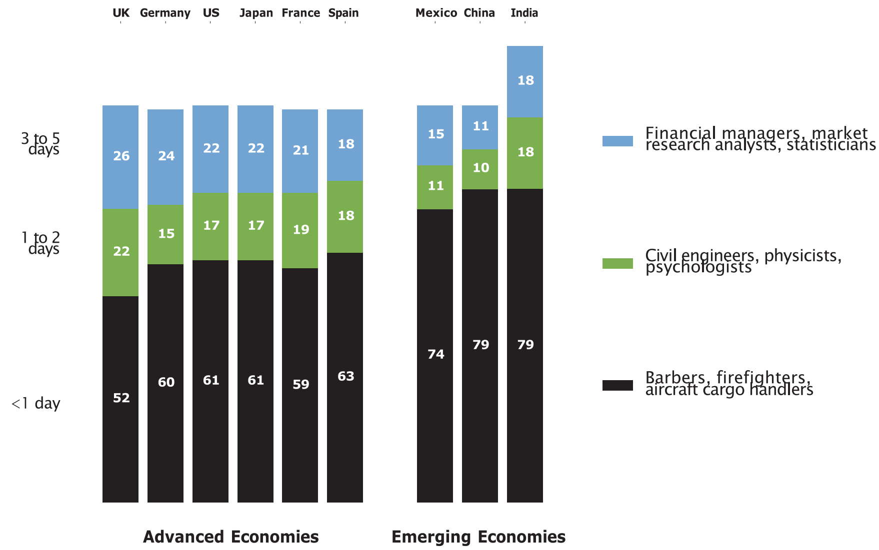

Fig. 2. Estimated remote-work potential across advanced and emerging economies, measured by the number of feasible remote working days per week based on occupational task composition [9]. The distribution highlights structural differences in digital readiness and reinforces the opportunity for remote AI workforce development in emerging regions.

3. How does one measure the effectiveness of remote training in AI/ML development, and what metrics indicate the success of such programs?
4. Considering the demands of industries for AI/ML expertise, how could remote training ensure that the workforce is equipped with the skills and knowledge necessary to meet these needs?

1.4. Scope

The scope of this paper includes the following:

- An introduction to the program design, including the program structure, recruitment process, curriculum and training plan, project work and the mentorship and leadership development plan.
- An analysis of the recruitment and participation trends, project impact, feedback and evaluation.
- The project internship outcomes and achievements including completed projects, publications, and graduate school or industry placements.
- Discussion on the challenges experienced, solutions and lessons learned.

2. Literature review

The adoption of virtual internships has garnered significant attention, particularly due to the constraints imposed by the COVID-19 pandemic. This section reviews the current literature on virtual internships. Research across various fields highlights the challenges and potential of virtual internships in enhancing accessibility and inclusivity in STEM education and other professional domains, while also emphasizing that quality and outcomes depend strongly on how technology mediates learning, feedback, and collaboration.

2.1. Remote internships & experiential learning

The shift to virtual and hybrid internships was initially a necessary response to the constraints imposed by the COVID-19 pandemic, but subsequent scholarship suggests that remote internships represent more than an emergency substitute for in-person placements. Rather, they constitute a distinct form of technology-mediated experiential learning. Hruska et al. examined the Smithsonian Environmental Research Center's 2020 Virtual Internship Program and demonstrated that intentionally designed virtual research experiences can broaden accessibility for students facing geographic, financial, or health-related barriers [7]. Their work emphasized structured project design, active mentoring, and deliberate community-building as central components of successful virtual programs. Building on this foundation, Hruska et al. further outline recommendations for normalizing virtual internships, proposing models that span a continuum from purely computer-based activities (e.g., digital data gathering and analysis) to hybrid hands-on research, emphasizing that design intentionality, not physical co-location, determines quality [7].

Similarly, Wong et al. found that work-from-home service-learning internships can enhance autonomy and motivation, yet require careful coordination, communication systems, and organizational support to sustain engagement and accountability [6]. These findings align with broader evidence that remote internships demand explicit infrastructural support. Bay et al. highlight the importance of structured workflows, clear communication channels, and logistical clarity in shaping professional development outcomes [13]. Feldman further argues that virtual internships can expand access for nontraditional learners, but only when institutions commit to sustained structural support [17]. In curricular contexts, Kernan and Basch demonstrate that transitioning a traditional semester-long internship to a virtual format requires explicit alignment with learning outcomes and structured components rather than simple modality substitution [18].

Additional studies reinforce that remote internships are embedded within broader socio-technical ecosystems. Paviotti et al. analyze how universities and employers adapted internships during the pandemic [10], while Deetlefs et al. show that behavioral interventions can increase participation in rural and remote internships [11]. Linkov et al.'s Contingency Internship Matrix for Epidemics (CIME) illustrates structured institutional responses to crisis-based internship disruptions [19], and Bowen highlights how boundary management and self-regulation shape remote work-integrated learning experiences [12]. Aleem et al.'s topic modeling analysis situates remote work within broader organizational and psychosocial trends [20]. Importantly, Lee cautions that virtual internships are not uniformly equitable; while many match in-person quality, disparities emerge across access, duration, and resource allocation [14].

Collectively, these studies suggest that virtual internships should be conceptualized not as simplified digital substitutes, but as structured experiential learning systems whose effectiveness depends on mentoring density, workflow clarity, technological mediation, and equitable resource distribution.

2.2. Computer supported collaborative learning (CSCL) in digital learning platforms

As remote internships usually unfold within digital environments, their learning processes are inherently shaped by platform-mediated interaction. From a CSCL perspective, remote internships can be understood as distributed learning systems in which shared artifacts, asynchronous communication, and tool-mediated feedback structure knowledge construction and professional identity formation.

Learning analytics research offers one perspective for examining these digitally mediated learning processes. Systematic reviews show that distance education increasingly relies on platform log data to analyze engagement patterns, communication behaviors, and performance trajectories [15]. In higher education, learning analytics has been positioned as a mechanism for moving from generic to personalized feedback, while raising important concerns around privacy, transparency, and algorithmic bias [16]. Similar tensions between personalization and risk are documented in K–12 education (primary and secondary school levels) [21].

Beyond analytics, adaptive learning technologies and tool-mediated feedback systems have demonstrated potential to support self-regulated learning when feedback is timely and tailored [22]. However, systematic reviews reveal that despite the availability of scalable feedback technologies, many educational contexts continue to rely on traditional instructor-dominated feedback structures [23]. Taken together, this body of research positions digital platforms not merely as delivery channels, but as active mediators of mentoring relationships, collaboration patterns, monitoring processes, and learning regulation in remote internship ecosystems.

2.3. AI workforce development & remote work

Recent advances in AI-assisted learning further reshape the landscape of remote experiential training. Systematic evidence indicates that AI-assisted feedback can enhance task-level, process-level, and self-regulatory learning outcomes when thoughtfully integrated [24]. Empirical studies show that students' engagement with generative AI tools is shaped by both perceived learning benefits and ethical considerations [25]. At the same time, assessment-focused work emphasizes the importance of balancing AI-assisted productivity with structured evaluation mechanisms to prevent overreliance and preserve foundational competence [26].

These developments are particularly consequential for AI/ML workforce preparation. Unlike many professional domains that require physical laboratories or localized equipment, AI training is inherently data-centric and computational. Learning artifacts, code repositories, model outputs, experiment logs, and documentation, are digital by default and can be shared, versioned, and evaluated remotely. Furthermore, access to cloud computing and shared computational infrastructure enables geographically distributed learners to conduct complex AI experiments independent of physical proximity. This location-independent data access and scalable compute availability make AI internships uniquely aligned with remote experiential models.

Experiential learning theory reinforces why such digitally mediated project work can remain effective when authenticity and learner engagement are preserved [27,28]. Integrating learner-generated data into project-based curricula further enhances ownership and engagement [29]. Accordingly, AI internships may be conceptualized as technology-mediated experiential ecosystems in which digital artifacts, platform-based collaboration, and AI-assisted scaffolding converge to support distributed workforce development.

In summary, virtual internships expand accessibility and flexibility [7,17], but require structured infrastructures and equitable resource allocation to ensure quality [13,14]. Learning analytics and AI-assisted feedback research demonstrate how digital tools can mediate mentoring, monitoring, and self-regulated learning in remote settings [15,16,21,22,24,26]. Continued theoretical and empirical investigation is therefore necessary to understand how remote internship ecosystems can equitably and effectively support AI workforce development [10,20].

3. Adopted program design and implementation

3.1. Design rationale and alignment with challenges in virtual internships

Previous studies on virtual internships have identified several recurring issues, including difficulties in monitoring progress, sustaining engagement and motivation, and ensuring effective communication across distributed teams. The literature also emphasizes challenges related to limited exposure to organizational culture, uneven learning opportunities, and the need for robust support infrastructures that can replace informal in-person mentoring and coordination.

The AI/ML remote internship program was designed to address these constraints through a structured, technology-mediated experiential learning model. In particular, the program emphasizes (i) clearly defined projects with staged deliverables, (ii) tool-supported coordination and feedback loops to improve visibility and accountability, and (iii) dedicated mentorship practices to sustain learning, collaboration, and professional growth in a remote setting. This section introduces the core elements of the program, from selection and onboarding to the six-month training structure, instructional tools, mentorship model, and project execution.

3.2. Program workflow and six-month structure

The program was designed as a six-month remote internship with an upfront selection phase that ensured baseline readiness for research-oriented AI/ML work. As illustrated in Fig. 3, the workflow includes two major phases: (1) the internship selection and onboarding process, and (2) the structured six-month internship experience with recurring coordination and feedback cycles.

Internship selection and onboarding: This phase was designed to identify candidates with the technical foundation and research readiness required for remote AI/ML project work. Candidates submit an application and proceeded through multiple assessments to evaluate programming proficiency, understanding of core ML concepts, and ability to read and synthesize scientific literature. Shortlisted candidates completed an interview focused on problem-solving and communication, after which successful applicants were onboarded into the internship.

Six-month internship experience: After onboarding, interns engaged in team-based projects with recurring meetings and milestone-based deliverables. Regular biweekly meetings provided structured opportunities to monitor progress, diagnose blockers, and deliver actionable feedback. In addition, interns participate in feedback sessions with explicit calls to action (CTA) and present interim and final outcomes to a guest audience, supporting iteration, accountability, and exposure to diverse perspectives. Fig. 4 visualizes the specific workflow and timeline of the 2024 Digital Twin cohort, tracking the interns' progression from initiation and skill development to system implementation. Furthermore, Table 1 provides a detailed and transparent breakdown of this six-month period, explicitly outlining the durations, specific activities, digital tools utilized, and assessment mechanisms for each phase to facilitate reproducibility.

Targeted skills development: The program embeded technical development within a broader set of professional competencies required for AI/ML careers. Interns received guided practice in communication,

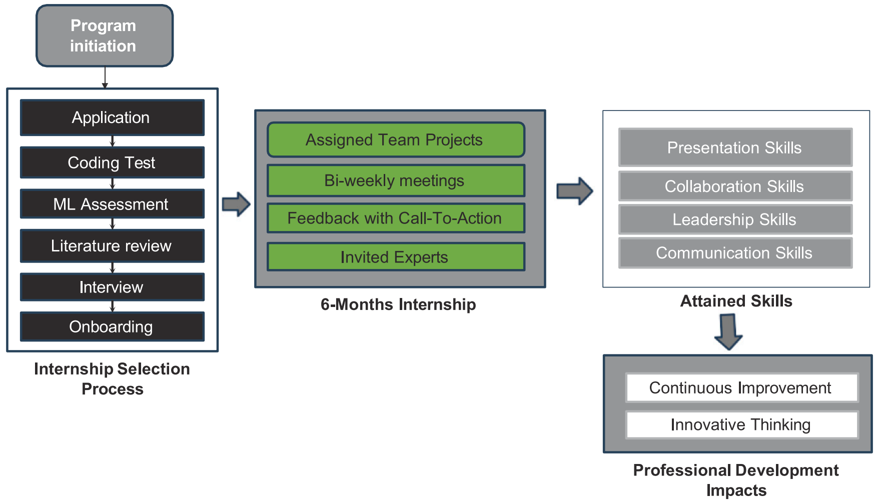

Fig. 3. Overall program workflow illustrating the structured progression from selection to the six-month internship phase. The diagram highlights iterative feedback cycles, mentorship checkpoints, and targeted technical and leadership skill development embedded throughout the program.

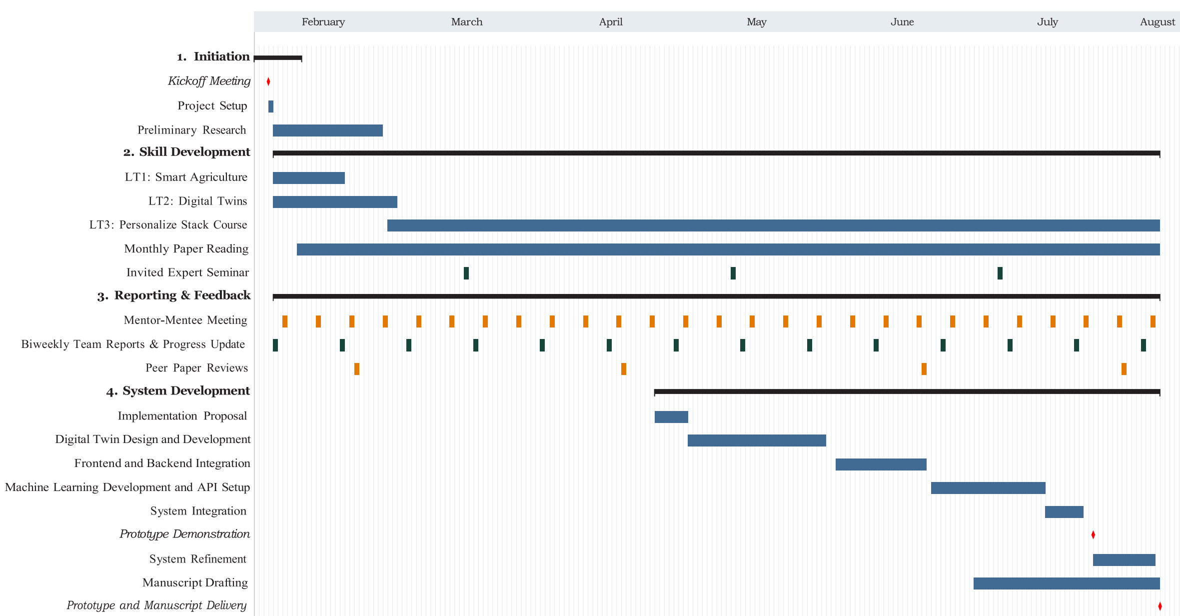

Fig. 4. Workflow and timeline of the 2024 Digital Twin internship cohort. The diagram illustrates the program's four structured phases: Initiation, Skill Development, Reporting & Feedback, and System Development. This visualizes the interns' progression from foundational domain learning (e.g., Smart Agriculture and Digital Twins) to continuous, mentor-guided system implementation, culminating in prototype demonstration and collaborative manuscript drafting.

scientific presentation, collaboration, and leadership. These competencies are reinforced through rotating responsibilities, continuous revision cycles, and repeated opportunities to communicate technical work to both technical and mixed audiences.

By combining staged selection, a time-bounded workflow, and recurring feedback loops, the program operationalized remote experiential learning in a way that supports both technical growth and professional formation.

3.3. Recruitment process

The recruitment process for the internship program was designed to identify highly qualified and motivated candidates who are passionate about AI and ML. The strategy encompasses targeted outreach, rigorous selection criteria, and a strong emphasis on diversity and inclusion, particularly in terms of gender representation and varied academic backgrounds, to ensure a rich and varied applicant pool.

Outreach and application: The program leverages multiple channels for outreach to attract a wide range of applicants. These channels include online job boards, professional networks, and social media platforms. Prospective candidates are required to submit an application that included their resume, a statement of purpose, and references. This comprehensive application process ensured the program received detailed information about each candidate's technical background and motivation.

Selection criteria: The selection process was rigorous and multifaceted, designed to evaluate candidates' technical abilities, critical thinking skills, and potential for growth. Applicants were assessed based on their academic performance, relevant experience, and demonstrated interest in AI/ML. The selection process included several stages:

- Coding test: To assess programming skills, candidates were required to complete a coding test that evaluated their proficiency in languages commonly used in AI/ML.

Table 1
Detailed six-month training breakdown (Based on the Digital Twin Cohort Timeline).

Phase	Duration & Sessions	Specific Activities	Tools Used	Learning Objective	Assessment Mechanism
1: Initiation	Weeks 1–4 (1 × 2-h meeting/ week)	Kickoff meeting, project setup, and preliminary domain literature review.	Slack, Google Meet, Trello	Establish research context and professional workflow habits.	Baseline research summary and environment setup.
2: Skill Development	Weeks 5–10 (Self-paced + 2 × 1.5-h seminars /month)	Learning tasks (Smart Agriculture, Digital Twins), monthly paper reading, and expert seminars.	Google Colab, Slack, Google Meet	Acquire foundational domain knowledge and software stack proficiency.	Completion of stack courses and paper summaries.
3: System Development	Weeks 11–18 (Daily async + 1 × 2-h biweekly review)	Digital Twin design, frontend/backend integration, and ML API setup.	GitHub, Colab, 3D Modeling Tools	Apply technical skills to build and deploy a functional prototype.	Biweekly team reports, code commits, prototype demo.
4: Refinement & Reporting	Weeks 19–24 (Daily async + 1 × 2-h biweekly review)	System refinement, mentor meetings, peer paper reviews, manuscript drafting.	Overleaf/LaTeX, Trello, Google Meet	Synthesize findings and practice collaborative scientific communication.	Final manuscript draft, peer-review feedback, final seminar.

- ML assessment: Candidates underwent an assessment that tested their understanding of key machine learning concepts and algorithms.
- Literature review: Applicants were required to analyze a piece of scientific literature, demonstrating their ability to comprehend and critically evaluate research.
- Interview: Successful candidates from the previous stages were invited for an interview, where they are assessed on their problem-solving abilities, communication skills, and cultural fit for the program.

3.4. Curriculum and training

The curriculum and training program was designed to equip interns with both technical expertise and essential soft skills required for success in the AI/ML field, with a particular focus on applications in the agricultural sector. The training is structured to ensure a comprehensive learning experience that combined theoretical knowledge with practical application.

- Core curriculum: The program offered a robust curriculum that allowed interns to explore foundational and advanced topics in AI and ML. This includes courses on data science principles, programming languages such as Python and R, and machine learning algorithms. Interns had access to online courses that provided a strong foundation in these areas, ensuring that all participants, regardless of their prior knowledge, could engage effectively with the material.
- Specialized training: Given the program's focus on the agricultural sector, specialized training is provided to interns, especially those who have a background in AI/ML but lack experience in agriculture. Online courses and workshops cover topics such as understanding greenhouse, sensor data analysis, predictive modeling for crop yields, and precision farming techniques. This targeted training ensures that all interns can apply their AI/ML skills to solve real-world agricultural problems effectively.
- Workshops and training sessions: the program places a strong emphasis on developing leadership skills. Interns participate in workshops that cover effective communication, teamwork, project management, and strategic thinking. These sessions are designed to help interns take on leadership roles within their teams and prepare them for future leadership positions in their careers.
- Presentation skills: To ensure that interns can effectively communicate their findings and ideas, the program includes training on presentation skills. This training covers how to create compelling visuals, structure presentations logically, and deliver messages clearly and confidently. Practice sessions and peer reviews are incorporated to help interns refine their presentation abilities.

Throughout the internship, interns are encouraged to engage in continuous learning and development. They are provided with resources and opportunities to explore new topics, attend webinars, and participate in online forums. This culture of continuous improvement helps interns stay current with the latest developments in AI/ML and agriculture.

3.5. Project-based learning and cohort implementation

The internship program allowed participants to engage in hands-on projects that applied their academic and technical skills to real-world challenges. Each project was chosen based on its relevance to current industry trends and its potential to provide impactful learning experiences for the interns. The project development process began with the advisory team assigning projects and providing the necessary datasets authentically curated to ensure their specificity to the project and reflection of real-world conditions. Throughout the project lifecycle, mentors guided interns and provided support during key stages such as background research, system development, and prototype presentation. The management team were also responsible for reviewing project deliverables to ensure they met the desired outcomes while providing feedback that can be used in the next iteration of the system development process.

Interns begin by conducting thorough background research and reviewing related works to identify the limitations and difficulties in existing methodologies. This helped them better understand the problems they were addressing and the available solutions to think up new approaches. After familiarizing themselves with the dataset, they move on to system development, iteratively refining their solutions and presenting prototypes for review. Throughout the process, they update their systems based on feedback and evolving project requirements until the final phase, where they are expected to document their work in manuscripts (geared toward peer-review) to summarize their contributions and results.

3.5.1. Cohort 1

In 2022, the inaugural cohort focused on the problem of optimizing temperature and humidity sensor placement in a greenhouse environment. The objective was to minimize the number of sensors required while preserving sufficient environmental coverage to ensure accurate monitoring and predictive modeling. This problem was selected to expose interns to a realistic trade-off scenario involving cost efficiency, spatial modeling, and data-driven decision-making in precision agriculture. The dataset consisted of weekly temperature readings collected from 56 IoT sensors deployed across the greenhouse over several months. Interns first conducted exploratory data analysis to understand spatial and temporal variability patterns, followed by feature engineering to identify critical environmental gradients. The project required interns to reason about spatial clustering, redundancy, and predictive stability under sensor reduction scenarios.

To address the optimization problem, multiple approaches were explored, including reinforcement learning, K-means++ clustering, and gradient boosting-based selection strategies [30–32]. The reinforcement learning formulation enabled interns to model sensor selection as a sequential decision-making process, while clustering-based approaches provided interpretable baselines for spatial coverage optimization. Gradient boosting methods were used to evaluate feature importance and identify high-informational-value sensor locations.

A key challenge encountered during this cohort was balancing model complexity with interpretability. Interns had to justify sensor reduction decisions not only in terms of predictive accuracy but also in practical deployment feasibility. Additionally, handling missing readings and temporal drift in environmental data required careful preprocessing and validation. This cohort served as a foundational training ground in experimental design, algorithm comparison, and research documentation. The experience informed subsequent cohorts by establishing a baseline methodology for sensor optimization, which was later extended through ensemble learning and digital twin simulations in later program iterations.

3.5.2. Cohort 2

In 2023, the project was divided into three teams, each focusing on a different challenge in precision agriculture and data analysis.

The first project, *Soil Quality Prediction* [33], explored the feasibility of using ML techniques to predict various physical and chemical properties of soil, such as moisture content, calcium (Ca), magnesium (Mg), potassium (K), effective cation exchange capacity (ECEC), and soil composition percentages (sand, silt, and clay) using digital soil images. This project aimed to develop a non-destructive soil analysis method by processing soil images through ML algorithms. The dataset provided for this project utilized a mobile phone camera for field sample collection, and further laboratory analysis on the acquired image samples resulted in additional structured data. This combined dataset became multimodal, involving both soil images and tabular physical and chemical properties data. The team compared different ML models and various image preprocessing techniques to determine the most suitable approach for predicting soil properties. They faced several challenges, including the variability in soil conditions, which impacted model consistency, and the complexity of accurately preprocessing the soil images for feature extraction. These challenges required the team to experiment with multiple data normalization methods and to fine-tune their preprocessing pipeline to ensure robust model performance. The results of this project demonstrated that ML can be effectively applied to soil analysis, offering a new approach that could help rural farmers estimate their soil quality for instances where traditional soil analysis machinery is unavailable due to location or cost.

The second project, *Optimal Sensor Placement using Ensemble Learning*, aimed to develop an ensemble of the approaches proposed by the first cohort. The ultimate goal was to maximize the efficiency and reliability of predictions for optimal sensors. To address this, they proposed using the sine cosine algorithm (SCA) and Particle swarm optimization (PSO), which improved the sensor placement strategy by identifying optimal sensor locations in real-time. The application of the SCA and PSO algorithms enabled the team to reduce the number of sensors while maximizing the accuracy and relevance of the data collected, improving the efficiency of the greenhouse monitoring system.

The third project, *Multimodal Livestock Feed Quality Prediction* [34], aimed to predict the nutritional composition of livestock feed by combining near-infrared (NIR) and red, green, and blue (RGB) images. Feed samples commonly used in livestock ration formulation were analyzed for their nutritional content using standard methods prescribed by the Association of Official Analytical Collaboration (AOAC). The interns worked to enhance classification techniques for multimodal images and adapt them for regression tasks to predict nutritional composition. Integrating data from different spectral sources was a significant challenge, which required careful preprocessing to align the NIR and RGB data. Despite these challenges, the team successfully developed an ML model capable of accurately predicting nutritional values. This non-destructive method for feed quality analysis can improve livestock feed formulation by enabling more precise and efficient assessments of feed nutritional content, ultimately enhancing feed quality and animal health.

3.5.3. Cohort 3

In 2024, a single project was assigned: *Optimal Sensor Placement in Strawberry Greenhouse using Digital Twin* [35]. This project focused on developing a digital twin of a strawberry greenhouse to simulate the optimal sensor deployment strategy. While the same dataset was used in the 2022, 2023, and 2024 sensor projects, the key difference in 2024 was incorporating a simulation component using digital twin technology. The project's complexity required a broader range of skills than previous projects, primarily focused on data analysis and ML. For this project, the team was divided into frontend, backend, and ML subteams to handle different aspects of the digital twin system. Leveraging 3D modeling tools and Game development engines, they developed a virtual replica of the greenhouse, enabling real-time visualization and interaction while replaying the provided IoT sensor data in real-time to simulate live conditions for the project accurately. A backend system managed this data, while ML algorithms were employed to analyze the data and optimize sensor placement. The combination of real-time data transfer, predictive modeling, and simulations allowed the team to determine the optimal number of sensors needed to capture critical environmental data, maximizing coverage while minimizing costs.

3.6. Mentorship and leadership development as a digital instructional strategy

The mentorship and leadership development component of the internship program was designed to cultivate both technical proficiency and essential leadership skills among the interns. The program was structured to provide a supportive learning environment while encouraging interns to take on significant leadership responsibilities.

- Mentorship program: Each intern was paired with a mentor who has expertise in the intern's specific area of focus. To ensure interns received high-quality guidance relevant to industry and academic standards, the program's advisory and mentorship team comprised experts with advanced degrees and specialized backgrounds in bio-industrial machinery, computer control, agricultural engineering, computer science, and materials engineering. Table 2 details the professional qualifications, disciplines, and institutional affiliations of the senior mentors and advisory team guiding the interns. These mentors provided ongoing guidance throughout the internship, helping interns navigate challenges, refine their project work, and develop their professional skills. The mentorship relationship was a cornerstone of the program, offering personalized support that is critical to the intern's success. Mentors assisted not only with technical aspects but also with career advice, fostering a holistic development experience.
  To ensure high-quality, responsive guidance and facilitate reproducibility, the dual mentorship model was operationalized with a low mentor-to-intern ratio of approximately 1:2 to 1:3. Prior to guiding the interns, both senior and peer mentors underwent a structured onboarding and preparation phase. This training aligned them with the program's pedagogical objectives and established strict feedback latency targets (e.g., maximum 24- to

Table 2
Mentor qualifications and affiliations.

Title	Name	Discipline/Qualifications	Affiliation
Dr.	Senior Mentor 1	PhD Bio-industrial Machinery Engineering	Michigan State University[a]
Dr.	Senior Mentor 2	PhD Agricultural and Biological Engineering	Shell[b]
Dr.	Senior Mentor 3	PhD IT Convergence Engineering	West Virginia University[a]
Dr.	Senior Mentor 4	PhD IT Convergence Engineering	University of Wyoming[a]
Prof.	Senior Mentor 5	PhD Computer Control and Automation	Kyungpook National University[c]
Prof.	Senior Mentor 6	PhD Biosystems Engineering	Kyungpook National University[c]
Mr.	Peer Mentor 1	PhD Student (Chemical and Biomolecular Engineering)	University of Nebraska-Lincoln[a]
Ms.	Peer Mentor 2	BSc. Computer Science	CLIMDES Lab[d]
Mr.	Peer Mentor 3	PhD Student (Biosystems & Agricultural Engineering)	Michigan State University[a]
Mr.	Peer Mentor 4	BSc. Metallurgical and Materials Engineering	CLIMDES Lab[d]

[a] Carnegie R1: Doctoral university with very high research activity.
[b] Global integrated energy company engaged in exploration, production, and energy technology development.
[c] National research university in South Korea.
[d] Research Group at Michigan State University.

48-hour response times on asynchronous channels). Mentors were also trained on standardizing their instructional use of digital tools, ensuring that feedback on Trello, GitHub, and Overleaf was consistent across teams. Peer mentors, typically successful alumni of previous cohorts, received targeted guidance on facilitating team dynamics, troubleshooting common coding blockers, and effectively escalating more complex theoretical or infrastructural issues to the senior mentors. Mentorship interactions were highly structured, combining biweekly synchronous milestone reviews with continuous, on-demand asynchronous support, ensuring interns never experienced progress-halting bottlenecks.

- Leadership training sessions: To prepare interns for future leadership roles, the program includes a series of leadership training sessions. These sessions covered various aspects of leadership, such as effective communication, decision-making, team management, and strategic thinking. Interns participated in workshops and discussions designed to build their confidence and competence in leading teams and managing projects.
- Rotational leadership: A unique feature of the program was the rotational leadership system. Each intern was given the opportunity to lead their team for a specified period. This rotation ensured that every intern gained hands-on experience in leadership, learning to make decisions, delegate tasks, and manage team dynamics. This experience is invaluable, as it prepares interns for leadership roles in their future careers. Within the rotational leadership structure, specific interns were designated as task leads for particular project components. This allowed them to take ownership of specific tasks, manage the execution of these tasks, and report on progress to the larger team. This role further enhanced their leadership experience, as they must balance managing a task with contributing to the overall project.
- Peer review and team-based presentations: Interns were encouraged to engage in peer review processes, providing feedback on each other's work. This collaborative approach enhanced the quality of the work produced and fostered a sense of accountability and teamwork. Additionally, interns were involved in team-based presentations, where they collectively presented their projects and findings. These presentations helped interns develop their public speaking and presentation skills, essential for any leadership role.
- Mentorship from previous Interns: The program also includes a peer mentorship component, where current interns receive guidance from former interns who have successfully completed the program. These former interns provide insights based on their experiences, helping the new interns avoid common pitfalls and maximize their learning. This continuity creates a strong support network within the program and reinforces a culture of mentorship and collaboration.

Overall, the program's mentorship and leadership development component was integral to the interns' growth, ensuring they leave the internship with advanced technical skills and the leadership capabilities necessary for future success.

3.7. Digital learning environment and tool-mediated instruction

To support remote experiential learning at scale, the internship adopted a digital learning environment in which collaboration and feedback were structured through a carefully selected suite of complementary tools. For instance, Trello was utilized for task coordination, while Slack facilitated asynchronous communication. Version control and peer code reviews were managed through GitHub, and Google Colab provided a shared computational workspace. Furthermore, the program integrated Overleaf and LaTeX for collaborative scientific writing, Google Meet for synchronous meetings, and GPT-based assistants for debugging and just-in-time support. Rather than treating these platforms as mere communication utilities, the program used them as instructional scaffolds and epistemic tools that externalized progress, preserved evidence of learning, and enabled feedback loops that are otherwise difficult to maintain in distributed settings.

Task visibility and accountability: *Trello* provided a shared representation of the project workflow by categorizing tasks into clear stages, such as to-do, in-progress, review, and done. This structure enabled interns and mentors to coordinate distributed work without relying on continuous supervision. As illustrated in Fig. 5, the boards facilitated task segmentation across teams and individuals, making ownership and deadlines explicit. This visibility reduced ambiguity around responsibilities, supported self-regulation, and allowed mentors to identify bottlenecks early. By making progress legible at a glance, Trello functioned as a lightweight mechanism for monitoring engagement and pacing across cohorts.

Low-bandwidth discussion and traceable decisions: *Slack* enabled threaded, asynchronous discussion that accommodated time zone differences and bandwidth constraints. Channels were organized by teams and/or tasks to keep questions, clarifications, and decisions in searchable threads. Compared to informal messaging tools commonly used in emerging regions (such as *WhatsApp*), this structured platform vastly improved research traceability: discussions about model choices, preprocessing decisions, and experiment outcomes remained archived and could be referenced during peer-review and supervisory feedback without being lost in a single continuous chat stream.

Version control as a learning artifact: *GitHub* was used to manage codebases, document iterative experimentation, and support review practices through issues, commits, and pull requests. This made development steps auditable and enabled mentors and peers to provide precise feedback on specific changes. In addition to supporting software engineering best practices, the repository history served as an evidentiary record of learning progress, including debugging milestones,

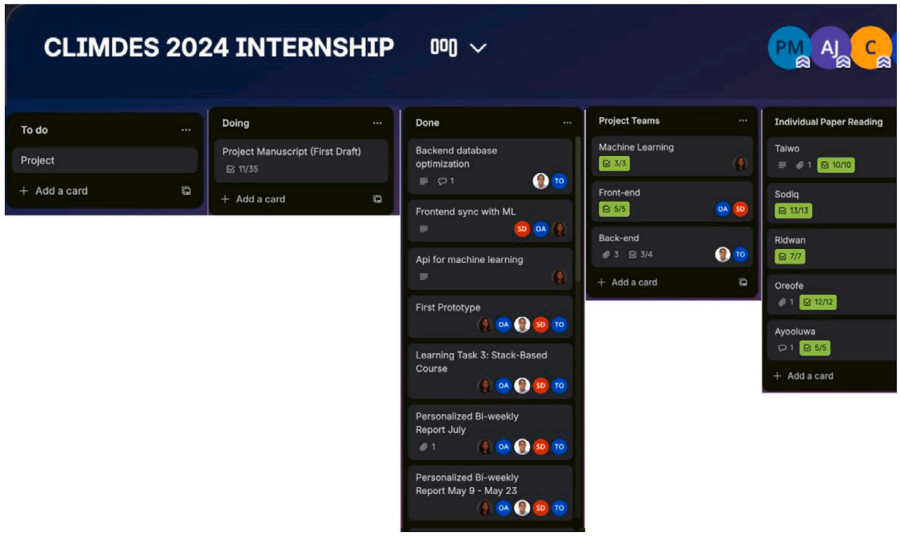

Fig. 5. Example Trello board from the 2024 Digital Twin cohort illustrating how task cards, workflow columns, and team segmentation functioned as collaborative cognition mediators. The board supported distributed project management across frontend, backend, and ML subteams while maintaining traceable progress and mentor oversight in a geographically dispersed setting.

refactoring decisions, and model iteration, thereby strengthening both accountability and reproducibility.

Shared compute and reproducible experimentation: *Google Colab* enabled collaborative coding and rapid experimentation through shared notebooks, reducing setup barriers for interns with heterogeneous hardware and software environments. Colab notebooks were used to standardize data loading and baseline experiments, facilitate mentor demonstration of workflows, and accelerate troubleshooting. This reduced overhead for mentors by minimizing environment-related issues and allowing feedback to focus on methodological reasoning rather than installation and configuration.

Collaborative scientific documentation: *Overleaf* supported collaborative manuscript writing and technical reporting, allowing interns to practice scientific communication while iterating on a shared document in parallel. Versioned writing and commenting features enabled mentors and peers to provide targeted feedback on clarity, structure, and evidence. Using *LaTeX* also aligned the internship outputs with research dissemination expectations and reinforced reproducible reporting habits, such as consistent figure referencing, table formatting, and citation management.

Synchronous coordination for milestone reviews: While most collaboration was asynchronous, *Google Meet* was used strategically for synchronous check-ins, milestone reviews, and presentations. These meetings anchored the internship cadence through activities like bi-weekly progress reviews, supported rapid clarification of complex issues, and provided practice in communicating technical work to varied audiences. To manage bandwidth constraints, meetings were prioritized for high-value interaction points, while routine updates remained asynchronous.

AI assistance with safeguards: The use of *GPT*-based assistants was encouraged primarily for debugging support, code explanation, and accelerating the comprehension of unfamiliar concepts, while explicit guidance discouraged overreliance. Interns were expected to validate suggestions through tests, documentation, and peer or mentor reviews. This positioning treated generative AI as a productivity aid within a supervised learning system rather than a substitute for foundational understanding. Peer-review and supervisory oversight provided an additional safeguard by checking that solutions aligned with accepted methodological practices and that reasoning remained transparent and defensible.

Tool-mediated feedback loops and mentor workload reduction: Across the internship, feedback was operationalized through tool-supported loops. Progress signals such as *Trello* boards, *GitHub* commits and issues, *Colab* outputs, and *Overleaf* revisions enabled mentors to review work asynchronously and provide targeted guidance with lower coordination costs. Peer review further distributed feedback responsibilities, allowing interns to learn from each other's implementations and writing while improving the speed of response to common issues. Collectively, this environment enabled remote mentorship at scale by turning day-to-day learning activities into visible, reviewable artifacts that supported timely, traceable, and action-oriented feedback.

4. Data analysis

This section provides an analysis of data spanning from the recruitment phase to the completion of the internship, focusing on trends in recruitment and participation, dropout rates, and overall success rates throughout the program. It also discusses the feedback and evaluation from previous participants.

4.1. Recruitment and participation trends

The recruitment process for the AI/ML remote internship program has evolved significantly across its three cohorts, reflecting broader shifts in the accessibility of remote internships and the growing demand for AI/ML skills. This subsection provides an analysis of recruitment strategies, applicant demographics, and participation trends observed during the internship period.

Recruitment strategies: The recruitment process leveraged various platforms, including academic networks, social media channels, and job boards, to attract many applicants. The program's remote nature broadened access to candidates from different regions and academic backgrounds, many of whom might not have been able to participate in traditional in-person internships. Fig. 6 illustrates the recruitment funnel, highlighting the number of applicants at each stage of the recruitment.

An analysis of the evaluation stages reveals an average drop-off rate of 18.08% for the 2023 cohort and a higher rate of 25.52% for the 2024 cohort. This increase in drop-off may be attributed to the more rigorous selection criteria and increased competition in the later cohort as the program gained prominence.

Similarly, Fig. 7 highlights the overall success rate, tracking the percentage of applicants who progressed from application to final selection across all three cohorts. The success rates indicate that 13.89% of applicants were selected in 2022, followed by 19.15% in 2023, and

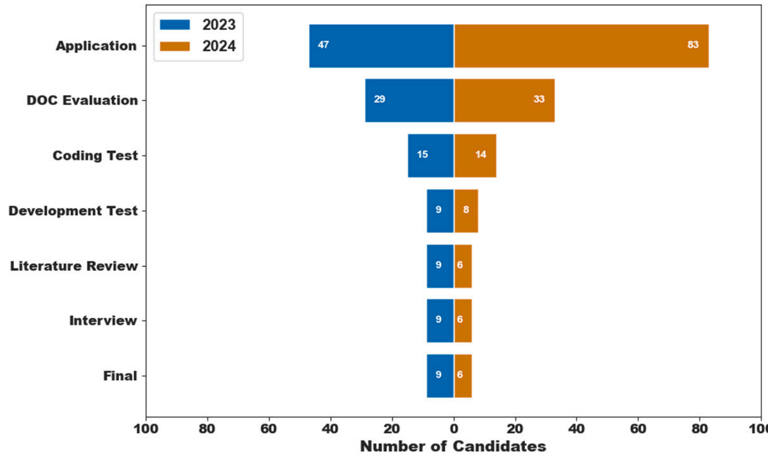

Fig. 6. Recruitment funnel for the 2023 and 2024 cohorts, showing progression from application submission to final selection. The funnel visualizes increasing applicant volume alongside selective filtering mechanisms designed to assess technical readiness and research aptitude.

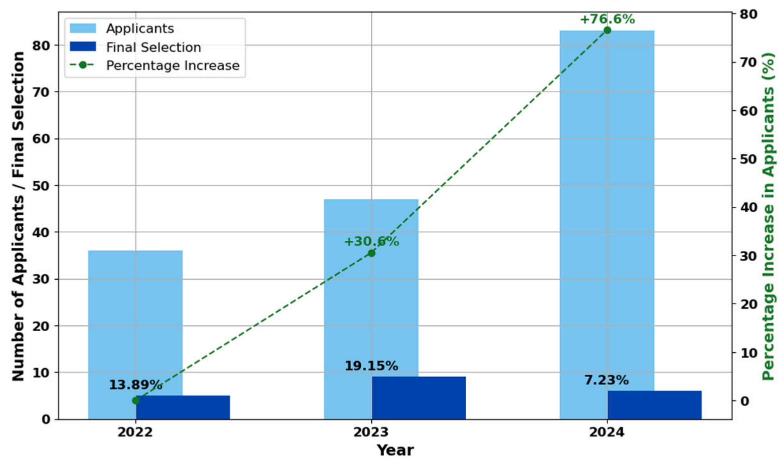

Fig. 7. Overall acceptance rate from application to final selection across cohorts. The variation across years reflects changes in applicant pool size, program visibility, and selection rigor.

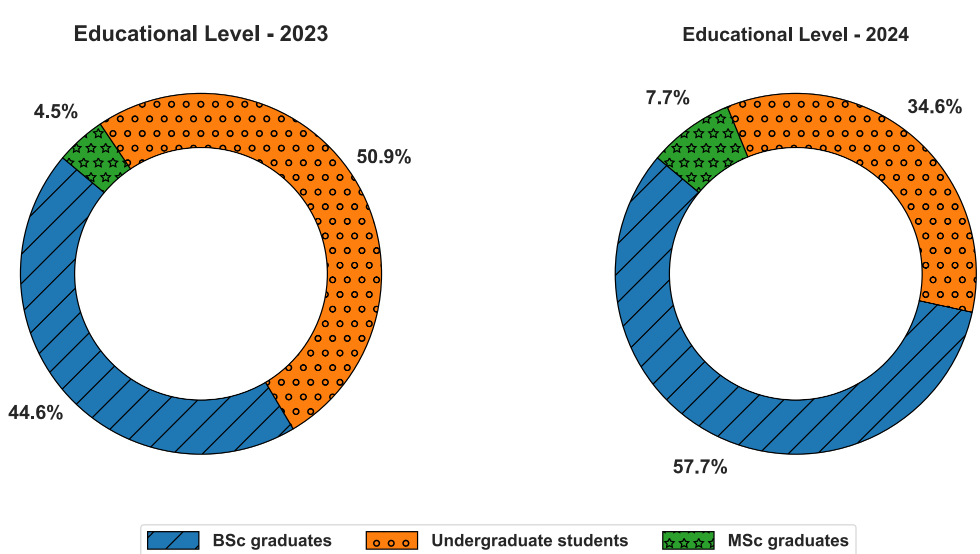

Fig. 8. Academic level distribution of applicants for the 2023 and 2024 cohorts. The diversity of undergraduate, graduate, and interdisciplinary backgrounds illustrates broad accessibility of the remote AI/ML internship model.

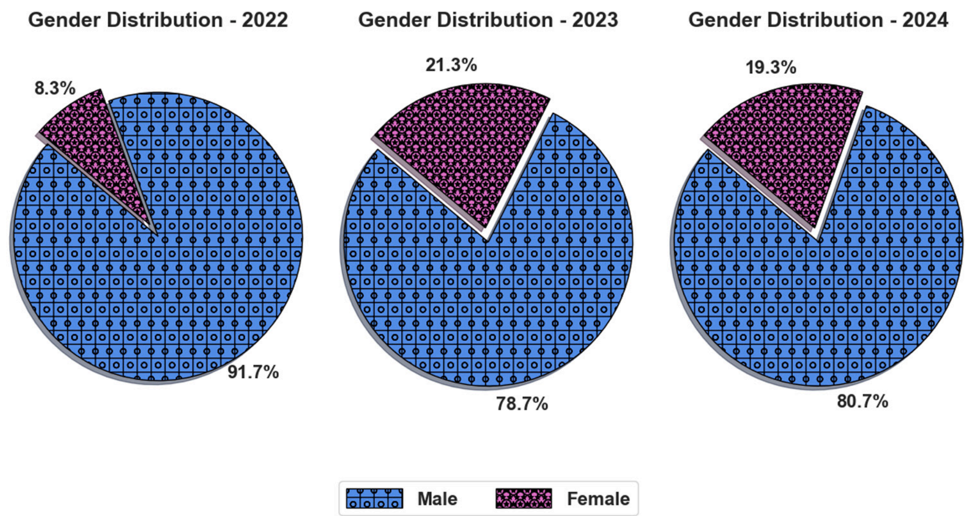

Fig. 9. Gender distribution of applicants across cohorts, demonstrating gradual increases in female participation over time and indicating improved outreach and accessibility in subsequent recruitment cycles.

7.23% in 2024. The noticeable dip in 2024 reflects both the increasing selectivity of the program and a larger applicant pool, highlighting the growing demand for participation as the program's reputation for high-quality training continues to expand. This internship program was able to attract a diverse range of academic levels, including undergraduates, recent graduates and Master's degree holders, as illustrated in Fig. 8.

Across the three cohorts, there was a noticeable increase in diversity, particularly in gender representation, with a steady rise in female applicants from 8.3% in the first cohort to 19.3% in the third cohort, as illustrated in Fig. 9. Across the three cohorts, there was also an increase in the number of applicants. As illustrated in Fig. 7, there was a 30.6% increase in applicants by the second cohort and a 76.6% increase by the third cohort. This highlighted the increased reach achieved with each recruitment. Applicants were evaluated based on their interest in research activities and agriculture. Notably, in the 2023 cohort, many applicants were primarily interested in agriculture, while the reverse was the case in the 2024 cohort, reflecting the program's growing visibility among broader research audiences.

Finally, the program consistently maintained an exceptionally high level of participation and retention. Across all cohorts, nearly all selected interns completed the full six-month duration (with only one dropout in the first cohort, none in the second, and one in the third). This low attrition rate can be attributed mostly to the program's strong mentorship support and structured accountability mechanisms as well as the selected participants' high level of motivation and dedication.

4.2. Feedback and program evaluation

Participants expressed high levels of satisfaction with the program, emphasizing its role in developing both technical skills and key interpersonal abilities such as communication, teamwork, and leadership. Interns appreciated the opportunity to engage in real-world projects that required both individual initiative and collaborative effort. Taking responsibility for leading teams and managing project deliverables enabled them to build essential leadership skills, while also enhancing their time management, adaptability, and communication abilities critical for success in fast-paced professional environments.

The perceived value of the program was evident in how participants described its impact on their technical growth. Many noted that the hands-on experience helped bridge the gap between theoretical knowledge and practical application, particularly in AI/ML and data science methodologies with real-world applications in agricultural technology. One of the key takeaways was their exposure to experimental methods in model development, which encouraged a shift from standard practices to more research-driven techniques.

Reflecting on the challenges of distributed work, one intern highlighted the value of the rotational leadership system in their feedback, noting: *"Serving as a team lead forced me to bridge the gap between just writing code and actually managing a distributed team's timeline across different time zones, which built my confidence immensely".*

Several interns shared defining moments that had a lasting impact on their personal and professional development. Overcoming technical challenges, such as optimizing models, interpreting complex research papers, or implementing advanced technological features, were often cited as pivotal experiences. The sense of accomplishment in solving these issues not only boosted their confidence but also reinforced their ability to approach real-world problems with innovative solutions.

Mentorship and peer collaboration were also highlighted as standout features of the program. Interns praised the supportive environment created by mentors, who provided both technical guidance and emotional support, helping them refine ideas, improve performance, and deliver high-quality project outcomes. The combination of mentorship and teamwork contributed significantly to the program's overall success. To systematically visualize this skill acquisition, Figs. 10 and 11 present competency heatmaps for the 2023 and 2024 cohorts, respectively. These heatmaps were generated directly from the interns' own feedback regarding the specific skills they built during the program. For the 2023 cohort (Fig. 10), self-reported skill development heavily clustered around foundational data analysis, machine learning, and project management. Conversely, reflecting the evolution and increased technical complexity of the program's curriculum, the feedback from the 2024 cohort (Fig. 11) highlights the acquisition of advanced, specialized capabilities such as digital twin architecture, cloud computing, and frontend engineering. Across both cohorts, however, core interpersonal competencies like leadership, collaboration, and confidence remained universally highlighted in the interns' feedback, highlighting the effectiveness of the dual mentorship model and the rotational leadership system in fostering holistic professional development.

4.3. Achievements

The program has achieved significant success in terms of post-internship outcomes. Over 60% of interns from the first and second cohorts transitioned to either graduate programs or secured full-time technical roles in private industry particularly in AI, data science, and

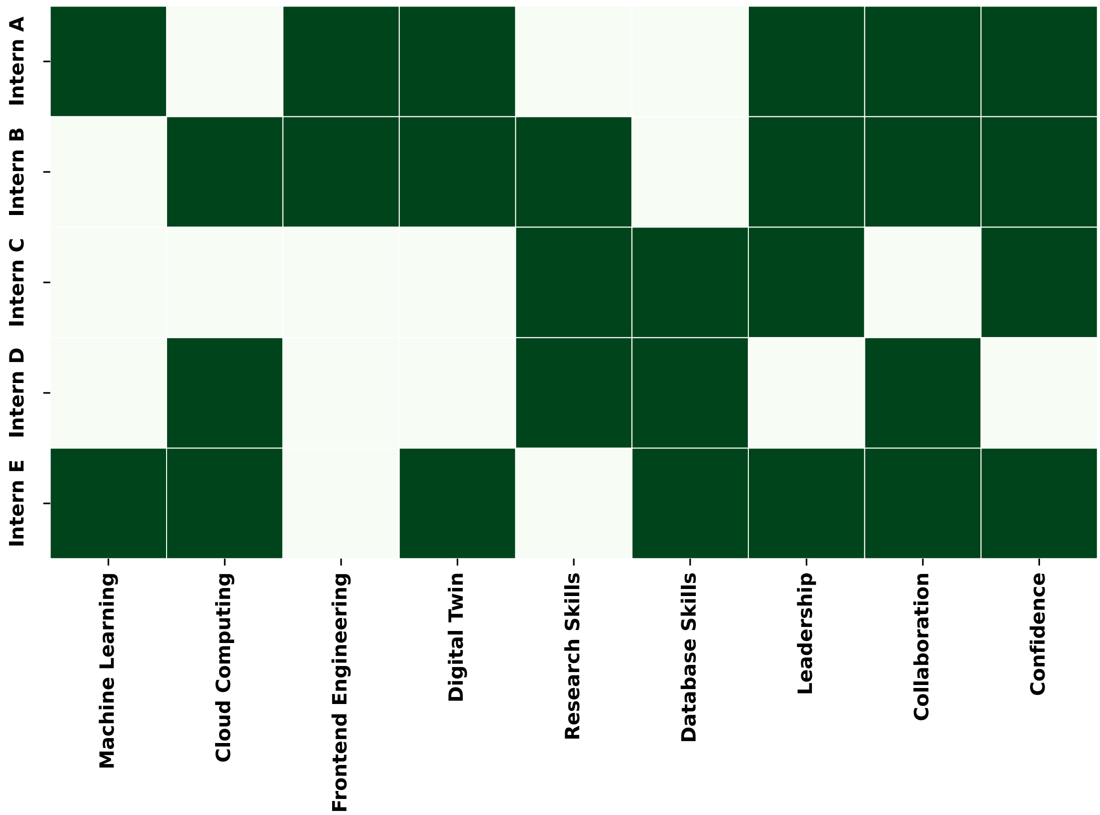

Fig. 10. Competency heatmap for the 2023 cohort (9 interns), derived from the interns' own feedback regarding the skills they built during the program. The map illustrates their self-reported development of technical skills (e.g., Machine Learning and Data Analysis) and professional competencies (e.g., Leadership and Adaptability).

related fields. Peer-reviewed articles co-authored by interns have been successfully published [33–36], highlighting the program's strong focus on research and publication as important indicators of success.

An analysis of individual intern outcomes reveals a strong correlation between active scientific dissemination and subsequent career advancement. Interns who engaged deeply with the program's collaborative scientific documentation tools to produce conference papers and poster presentations (such as those affiliated with the CLIMDES Lab) were disproportionately successful in securing subsequent graduate school placements and formal research assistance roles. This suggests that the program's emphasis on treating software version control and collaborative writing as "epistemic tools" directly translates into tangible academic currency for the participants.

Table 3 provides a detailed cohort-level breakdown of these achievements, documenting completion rates, peer-reviewed publications, conference attendance, and career progression across all three cohorts of the program. Furthermore, the "Research Assistance/Employment" metric in Table 3 is directly validated by the program's internal retention pipeline. As evidenced previously in Table 2, several high-performing alumni (such as those affiliated with the CLIMDES Lab) were subsequently hired as research assistants and integrated back into the program's advisory team as peer mentors. These outcomes highlight the program's effectiveness in preparing participants for both academic and professional advancement.

5. Discussion

5.1. Challenges and solutions

Throughout the internship's cycles, several challenges were encountered, most notably related to time zone differences and balancing internship commitments with academic or professional responsibilities. This issue became particularly evident during the second cohort, as many interns returned to in-person academic programs post-pandemic. The conflicting schedules created difficulties in coordinating team meetings and completing project milestones.

To address these challenges, more flexible meeting schedules were introduced, along with staggered deadlines to accommodate interns in different time zones and with varying academic or work obligations. Additionally, communication channels were optimized, moving away from bandwidth-heavy video calls toward asynchronous tools such as Slack and Google Docs to overcome connectivity issues, particularly for interns in regions with limited internet bandwidth.

Another challenge involved delays in feedback loops, where interns required timely responses from mentors to make progress. To mitigate this, the program introduced dual mentorship, pairing each intern with both a senior mentor and a peer mentor (often a previous intern), to ensure quicker turnaround on queries while also providing a support network that fosters inclusivity and shared learning.

5.2. Lessons learned

A key lesson from managing the internship program is the importance of adaptability in both structure and delivery. The need for flexibility, whether in project timelines, meeting schedules, or communication methods, was crucial in keeping interns engaged and on track, especially as personal and professional obligations evolved. This flexibility will remain a core feature in future iterations of the program.

Furthermore, the mentorship model proved to be a pivotal element in enhancing the learning experience. Pairing interns with peer mentors allowed for a more accessible and relatable form of guidance, while

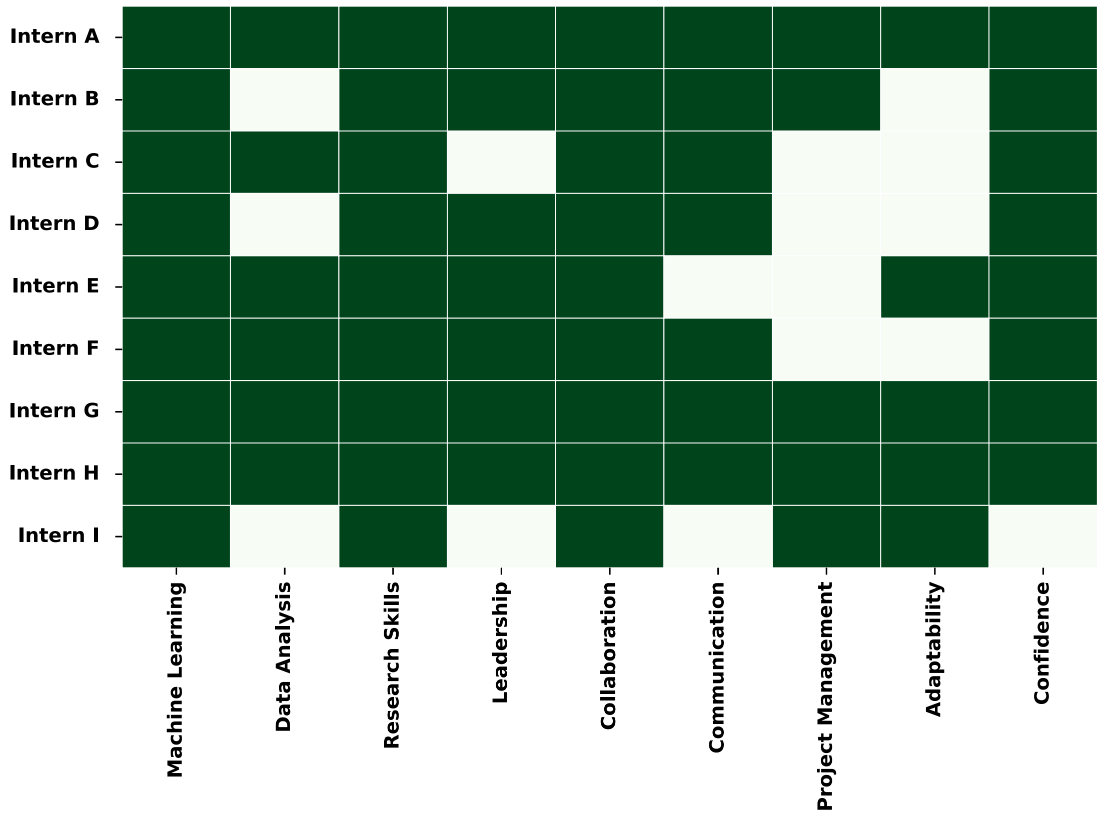

Fig. 11. Competency heatmap for the 2024 cohort (5 interns), based directly on interns' feedback about the skills they built. It highlights their self-reported shift toward highly specialized technical skills (e.g., Digital Twin, Frontend Engineering, Cloud Computing) alongside core professional competencies, reflecting the increased complexity of the project.

Table 3
Post-internship outcomes and achievements (Cohorts 1-3).

Metric/Outcome	Cohort 1 (2022)	Cohort 2 (2023)	Cohort 3 (2024)
Total Interns Completed	5	9	5
Peer-Reviewed Publications	100% (5/5)	89% (8/9)	100% (5/5)
Conference Attendance	100% (5/5)	77% (7/9)	60% (3/5)
Graduate School Progression	80% (4/5)	22% (2/9)	20% (1/5)
Research Assistance/Employment	100% (5/5)	33% (3/9)	40% (2/5)

Two of the participants who progressed to graduate school also held research assistant positions.

industry mentors provided technical expertise and professional development insights. Moving forward, formalizing this dual mentorship approach will be a priority.

A final takeaway is the significant role of real-world projects. Interns repeatedly emphasized that working on meaningful, real-world problems aligned with industry needs not only improved their technical skills but also gave them confidence in their ability to contribute to professional projects post-internship. As a result, future iterations will continue to prioritize projects that provide both practical and strategic value to interns and participating organizations.

Finally, tracking the program's evolution highlights the importance of iterative pedagogical improvements. Table 4 outlines the specific structural design changes made across the cohorts to address emerging challenges and increasing project complexity.

5.3. Comparisons to similar programs

Compared to other virtual internship programs, our initiative stands out in several key areas, though it also shares common ground with other successful programs in important ways. One prominent challenge identified in other programs is the delayed mentorship feedback, which can lead to feelings of isolation and disconnection for interns [37]. Similar challenges were highlighted in the outlined rules for managing virtual internships [38], which stress the importance of regular, structured communication to prevent disengagement and isolation in virtual settings. Our program has successfully countered this by implementing a structured feedback system where interns receive both technical guidance and peer support through dual mentorship, fostering a collaborative and inclusive learning environment. Table 5 summarizes these comparative evaluation metrics, contrasting the structural features of our remote AI/ML program against common trends and challenges identified in recent virtual internship literature.

In terms of project involvement, many internship programs fall short in giving interns access to relevant resources and meaningful projects. For example, the implementation by [37] highlighted that while some interns were given access to learning platforms to up-skill, others had restricted access to these resources, limiting their ability to fully engage in relevant project work. Interns in our program receive full and equal access to databases, project repositories, and training resources, ensuring that they are well-equipped to tackle industry-relevant challenges. In a study by Zahariev et al. [39], only 32% of interns selected ICT-related projects, not due to program limitations

Table 4
Iterative design changes and structural improvements across cohorts.

Cohort Transition	Pedagogical Challenge/Observation	Implemented design change
Cohort 1 to 2	A single-team focus limited the diversity of ML techniques explored and reduced peer-to-peer cross-pollination.	Transitioned to multi-team parallel projects. Introduced ensemble learning methodologies to synthesize multiple approaches for higher predictive accuracy.
Cohort 2 to 3	Transitioning to Digital Twin technology significantly increased the complexity of the software engineering required beyond standard ML tasks.	Segmented teams into highly specialized roles (frontend, backend, and ML) to manage distributed cognitive load and mirror industry software development lifecycles.

Table 5
Comparison of virtual internship programs by evaluation metrics.

Evaluation metric	Trends in similar programs (Literature)	Our Remote AI/ML program
Mentorship & Feedback	Prone to delayed feedback, leading to isolation and disengagement [37,38].	Structured dual mentorship (senior and peer mentors) ensuring rapid, continuous feedback.
Resource Access	Often features restricted or unequal access to learning platforms and tools [37].	Full, equal access to high-compute digital environments, datasets, and project repositories.
Project Focus	Low engagement in tech tasks (e.g., only 32% selecting ICT projects) [39].	100% focus on cutting-edge ICT skill development (AI/ML in precision agriculture).
Peer Interaction	Limited peer interaction is frequently cited as a major programmatic drawback [37].	Structured peer-paper reviews, collaborative code development, and independent inter-team sessions.
Schedule Flexibility	Recognized as valuable but sometimes constrained by synchronous routines [37,38].	Highly flexible, asynchronous tool-mediated workflows to accommodate global time zones.
Post-Internship Support	Often functions solely as a short-term learning experience [40].	Long-term focus with ongoing mentorship, publication co-authorship, and graduate placement support.

but rather intern choice, which narrowed the scope of their learning. This points to a trend in some programs where access to cutting-edge technology and real-world applications is limited, even when interns have the choice to pursue such projects. By contrast, our program is fully focused on ICT skill development, with a practical emphasis on AI/ML applications in agriculture. This ensures that even though the application domain is conventionally non-ICT related, interns are set on the path to acquiring advanced skills that are then applied to aid in sustainable farming practices, thus contributing to critical solutions in the agri-tech space.

Collaboration is another area where our program excels. While limited peer interaction is often cited as a drawback in virtual internships [37], we address this through structured peer-paper review sessions, collaborative team projects, and inter-team sessions, where interns work independently without direct management oversight. The discussions and debates that emerge from the peer-paper review sessions often lead to ongoing one-on-one or group conversations, where participants explore the intricacies of the reviewed papers in greater depth. These exchanges frequently result in the development of potential implementation strategies and provide valuable feedback on the sessions' effectiveness. This not only promotes peer-to-peer learning but also helps interns build self-reliance and collaborative skills in a team setting, enhancing their ability to work effectively without constant supervision.

Moreover, time management flexibility is a standout feature of our program, similar to the implementation [37], where interns appreciated the flexible working hours that allowed them to take breaks as needed and work according to their schedules. This flexibility is essential in virtual internships to enable participants to balance their workload with other commitments while avoiding the rigid routines associated with traditional office hours. This also positively influences learning and performance [38], as it effectively accommodates diverse needs, particularly for interns managing academic or part-time job commitments.

Finally, our program's focus on long-term professional growth sets it apart. Opportunities such as graduate school scholarships for top performers and ongoing post-internship mentorship foster continuous career development. This approach, as highlighted by Ezeafulukwe et al. [40], ensures that the program serves as more than just a learning experience and acts as a launchpad for future success, equipping interns with the skills and networks needed for their professional journey.

5.4. Limitations

While the remote AI/ML internship program demonstrated high success and retention rates, it is important to acknowledge certain limitations, particularly regarding selection bias. As detailed in our methodology, applicants underwent a rigorous four-step screening phase encompassing a coding test, a machine learning assessment, a literature review, and a final interview. Consequently, the program selected exceptionally motivated, technically prepared, and research-oriented candidates, accepting only a small fraction of the total applicant pool (e.g., a 7.23% acceptance rate in the 2024 cohort).

Because the program naturally filtered for the best potential interns, the strong post-program outcomes, lack of drop-offs, and high publication rates may be partly attributed to the pre-existing aptitude and dedication of the participants rather than solely to the program's instructional and mentorship design. While the structured digital environment and dual mentorship model provided necessary scaffolds, future research should evaluate how this pedagogical framework performs with a broader, less selectively filtered cohort to fully untangle the program's intrinsic impact from baseline participant readiness.

6. Conclusion and future directions

6.1. Conclusions

This paper presents the development, implementation, and outcomes of a remote internship program designed to build capacity in the future workforce for artificial intelligence and machine learning (AI/ML) through hands-on, real-world projects. The program addressed the growing demand for accessible remote learning and workforce development opportunities following the COVID-19 pandemic by integrating structured mentorship, leadership development, and technical skill training within an applied research environment. Several factors contributed to the success of the program, including its adaptability to evolving intern needs, a strong dual-mentorship framework, and an emphasis on practical projects aligned with real-world industry and research challenges. Flexible scheduling, collaborative learning, inclusivity, and peer-to-peer engagement further strengthened participant experience and retention. Program outcomes demonstrate substantial impact, with more than 60% of participants progressing to graduate studies or securing technical positions in industry. In addition, multiple peer-reviewed publications co-authored by interns highlight the program's success in bridging theoretical knowledge with applied problem-solving, preparing participants for both academic advancement and professional careers in AI-enabled systems.

6.2. Future plans

Looking ahead, the evolution of the internship program will pivot from establishing a baseline remote structure to investigating scalable, technology-enhanced pedagogical models. Based on the challenges and outcomes observed across the first three cohorts, future iterations will focus specifically on addressing resource disparities and scaling mentorship capacity through artificial intelligence.

Firstly, this study highlighted that while remote infrastructure can successfully deliver training, significant resource constraints remain regarding data acquisition. Specifically, in the area of agriculture, the high costs and specialized expertise required to collect high-quality datasets (such as multimodal soil images and IoT sensor data) pose a continuous barrier to advanced AI research in emerging economies. To overcome this limitation in future cohorts, the program will formally establish cross-continental research partnerships, bridging data-rich institutions in regions such as the United States with remote learners in Africa. This collaborative framework will pool together data collection resources and domain expertise to help ensure that interns distributed across diverse geographies have equitable and sustainable access to cutting-edge and authentic datasets for their projects without the prohibitive localized costs.

Secondly, as the program scales to accommodate more participants, human mentorship capacity presents a critical bottleneck. Because senior mentors manage these responsibilities alongside full-time academic or industry roles, minimizing routine troubleshooting and reducing meeting times is essential. To address this, future internships will be launched as structured pilots to test the integration of Generative AI as a pedagogical "co-researcher". By utilizing AI-assisted debugging, code explanation, and literature synthesis, the program will aim to handle baseline inquiries automatically. This intervention is expected to significantly reduce mentor feedback latency, allowing human experts to focus their limited time exclusively on high-level methodological reasoning and developing research strategies.

To ensure the responsible implementation of this AI-driven mentorship system, the pilot will enforce strict data privacy protocols by anonymizing intern codes and inputs. Furthermore, advisory mentors will routinely audit the AI-generated guidance to mitigate algorithmic bias and ensure the tool provides equitable, culturally responsive, and methodologically sound support to all participants.

CRediT authorship contribution statement

Judith Nkechinyere Njoku: Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation. **Patience Chizoba Mba:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Senorpe Asem-Hiablie:** Writing – review & editing, Supervision, Investigation, Formal analysis, Conceptualization. **Daniel Dooyum Uyeh:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Acknowledgments

The authors would like to express their sincere gratitude to Prof. Tusan Park, Prof. Rammohan Mallipeddi, and Prof. Yu Shin Ha for their foundational contributions to this remote internship effort. As members of the original founding team, their early ideas, strategic conceptualization, and dedicated efforts during the initial design phases were instrumental in establishing the program's technology-mediated framework and ensuring its long-term success.

Funding information

This work was supported by Michigan State University's internal funds.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article.

References

[1] Kristen Reid CC, Potter ADJ. Virtual internships in open and distance learning contexts: Improving access, participation, and success for underrepresented students. Distance Educ 2023;44(2):267–83. http://dx.doi.org/10.1080/01587919.2023.2209029.

[2] Erb M. Fortune 500 executives tell us what their post-COVID workplaces will look like. 2021, Great Place To Work® URL https://www.greatplacetowork.com/resources/blog/new-survey-fortune-500-ceos-reveal-what-the-future-office-will-look-like-post-covid.

[3] Gartner I. Gartner HR survey reveals 88% of organizations have encouraged or required employees to work from home due to coronavirus. 2020, URL https://www.gartner.com/en/newsroom/press-releases/2020-03-19-gartner-hr-survey-reveals-88–of-organizations-have-e. (Accessed: 05 September 2024).

[4] Dalton M, Groen JA. Telework during the COVID-19 pandemic: estimates using the 2021 business response survey. Mon Labor Rev 2022. URL https://www.bls.gov/opub/mlr/2022/article/telework-during-the-covid-19-pandemic.htm.

[5] UNESCO. Education: From COVID-19 school closures to recovery. 2022, United Nations Educational, Scientific and Cultural Organization. URL https://www.unesco.org/en/covid-19/education-response.

[6] Wong MML, Lau KH, Chan CWF. The impacts and success factors of a work-from-home service-learning internship during COVID-19. J Work-Applied Manag 2021;13(2):284–301. http://dx.doi.org/10.1108/JWAM-01-2021-0003.

[7] Hruska AM, Cawood A, Pagenkopp Lohan KM, Ogburn MB, Komatsu KJ. Going remote: Recommendations for normalizing virtual internships. Ecosphere 2022;13(3):e3961. http://dx.doi.org/10.1002/ecs2.3961.

[8] Li C, Lalani F. The rise of online learning during the COVID-19 pandemic. In: World economic forum. 2020, p. 2.

[9] Company M. The future of remote work: An analysis of 2,000 tasks, 800 jobs, and 9 countries. 2020, URL https://www.mckinsey.com/featured-insights/future-of-work/whats-next-for-remote-work-an-analysis-of-2000-tasks-800-jobs-and-nine-countries.

[10] Paviotti G, Aleffi C, Ferrara A. Toward remote work: Online internship in time of pandemics. Excel Innov Learn Teach 2023;1(2023):70–82. http://dx.doi.org/10.3280/exioa1-2023oa16045.

[11] Deetlefs AJ, Chalmers J, Tindall K, Wiryakusuma-McLeod C, Bennett S, Hay I, Humphries J, Eady MJ, Cronin L, Rudd K. Applying behavioral insights to increase rural and remote internships: Results from two randomized controlled trials. J Behav Exp Econ 2021;92:101664. http://dx.doi.org/10.1016/j.socec.2021.101664.

[12] Bowen T. Work-integrated learning placements and remote working: Experiential learning online.. Int J Work-Integrated Learn 2020;21(4):377–86.

[13] Bay J, Fillenwarth GM, Masters-Wheeler C. Remote TPC internships: Infrastructures for success. J Tech Writ Commun 2022;52(4):386–410. http://dx.doi.org/10.1177/00472816211041315.

[14] Lee C. Virtual internships as alternative work-based learning: Examining access, quality, and outcomes for underserved students. Comput Educ 2025;239:105439. http://dx.doi.org/10.1016/j.compedu.2025.105439, URL https://www.sciencedirect.com/science/article/pii/S0360131525002076.

[15] Palanci A, Yılmaz RM, Turan Z. Learning analytics in distance education: A systematic review study. Educ Inf Technol 2024;29(17):22629–50. http://dx.doi.org/10.1007/s10639-024-12737-5.

[16] Sharif H, Atif A. The evolving classroom: How learning analytics is shaping the future of education and feedback mechanisms. Educ Sci 2024;14(2). http://dx.doi.org/10.3390/educsci14020176, URL https://www.mdpi.com/2227-7102/14/2/176.

[17] Feldman E. Virtual internships during the COVID–19 pandemic and beyond. New Horizons Adult Educ Hum Resour Dev 2021;33(2):46–51. http://dx.doi.org/10.1002/nha3.20314.

[18] Kernan WD, Basch CH. Adaptation of an in-person internship to a virtual format for public health undergraduates. JMIR Public Health Surveill 2022;8(3):e35252. http://dx.doi.org/10.2196/35252.

[19] Linkov F, Khanijahani A, Swanson-Bearman B, Akinci F. Remote internship: Practical approaches to sustaining student internships amid public health epidemics. J Health Adm Educ 2021;38(1):363–76.

[20] Aleem M, Sufyan M, Ameer I, Mustak M. Remote work and the COVID-19 pandemic: An artificial intelligence-based topic modeling and a future agenda. J Bus Res 2023;154:113303. http://dx.doi.org/10.1016/j.jbusres.2022.113303.

[21] Paolucci C, Vancini S, Bex II RT, Cavanaugh C, Salama C, de Araujo Z. A review of learning analytics opportunities and challenges for K-12 education. Heliyon 2024;10(4):e25767. http://dx.doi.org/10.1016/j.heliyon.2024.e25767, URL https://www.sciencedirect.com/science/article/pii/S2405844024017985.

[22] Mejeh M, Sarbach L, Hascher T. Effects of adaptive feedback through a digital tool: A mixed-methods study on the course of self-regulated learning. Educ Inf Technol 2024;29(14):1–43. http://dx.doi.org/10.1007/s10639-024-12510-8.

[23] Okumu M, Lammert C, Gokmen C, Mule F. Mapping the feedback landscape: a systematic review of research on feedback sources, methods, and technologies in preservice teacher education. Front Educ 2025;Volume 10 - 2025. http://dx.doi.org/10.3389/feduc.2025.1657737, URL https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2025.1657737.

[24] Ba S, Yang L, Yan Z, Looi CK, Gašević D. Unraveling the mechanisms and effectiveness of AI-assisted feedback in education: A systematic literature review. Comput Educ Open 2025;9:100284. http://dx.doi.org/10.1016/j.caeo.2025.100284, URL https://www.sciencedirect.com/science/article/pii/S2666557325000436.

[25] Heyer C, Nilsson EM, Pedersen J. AI-assisted learning in HCI education: Opportunities and dilemmas from a student perspective. In: Proceedings of the 7th annual symposium on HCI education. EduCHI '25, New York, NY, USA: Association for Computing Machinery; 2025, http://dx.doi.org/10.1145/3742901.3742910, URL https://doi.org/10.1145/3742901.3742910.

[26] Elshall AS, Badir A. Balancing AI-assisted learning and traditional assessment: the fact assessment in environmental data science education. Front Educ 2025;Volume 10 - 2025. http://dx.doi.org/10.3389/feduc.2025.1596462, URL https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2025.1596462.

[27] Gross Z, Rutland SD. Experiential learning in informal educational settings. Int Rev Educ 2017;63(1):1–8. http://dx.doi.org/10.1007/s11159-017-9625-6.

[28] Kong Y. The role of experiential learning on students' motivation and classroom engagement. Front Psychol 2021;Volume 12 - 2021. http://dx.doi.org/10.3389/fpsyg.2021.771272, URL https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.771272.

[29] Li S-TT, Abramson E, Hilgenberg S, Lichtenstein C, Lockspeiser T. Enhancing learner engagement through experiential learning with learner-generated data. Acad Pediatr 2023;23(4):846–8. http://dx.doi.org/10.1016/j.acap.2022.11.002, URL https://www.sciencedirect.com/science/article/pii/S1876285922005617.

[30] Uyeh DD, Akinsoji A, Mallipeddi R, Asem-Hiablie S, Amaizu M, Ha Y, Park T. A reinforcement learning approach for optimal placement of sensors in protected cultivation systems. IEEE Access 2021;9:100781–800. http://dx.doi.org/10.1109/ACCESS.2021.3096828.

[31] Uyeh DD, Olayinka I, Mallipeddi R, Asem-Hiablie S, Amaizu M, Ha Y, Park T. Grid search for lowest root mean squared error in predicting optimal sensor location in protected cultivation systems. Front Plant Sci 2022;13. http://dx.doi.org/10.3389/fpls.2022.920284.

[32] Uyeh DD, Akinsoji A, Asem-Hiablie S, Bassey BI, Osinuga A, Mallipeddi R, Amaizu M, Ha Y, Park T. An online machine learning-based sensors clustering system for efficient and cost-effective environmental monitoring in controlled environment agriculture. Comput Electron Agric 2022;199:107139. http://dx.doi.org/10.1016/j.compag.2022.107139.

[33] Mba PC, Busoye OM, Ajayi JT, Njoku JN, Nwakanma CI, Asem-Hiablie S, Mallipeddi R, Park T, Uyeh DD. Estimation of physico-chemical properties of soil using machine learning. Smart Agric Technol 2024;9:100679. http://dx.doi.org/10.1016/j.atech.2024.100679.

[34] Oluwabukunmi OB, Joseph AA, Udurume M, Njoku JN, Nwakanma CI, Asem-Hiablie S, Mallipeddi R, Park T, Uyeh DD. Advanced multimodal prediction of components of livestock feed materials using knowledge distillation. IEEE Trans AgriFood Electron 2025. http://dx.doi.org/10.1109/TAFE.2025.3548949.

[35] Akintan O, Babawale S, Olomo A, Adeyemo R, Opadotun O, Ajayi JT, Mba PC, Njoku JN, Chesang A, Zahid A, et al. A digital twin framework for sensor selection and microclimate monitoring in greenhouses. AgriEngineering 2025;7(10):315. http://dx.doi.org/10.3390/agriengineering7100315.

[36] Ajani OS, Aboyeji E, Mallipeddi R, Uyeh DD, Ha Y, Park T. A genetic programming-based optimal sensor placement for greenhouse monitoring and control. Front Plant Sci 2023;14. http://dx.doi.org/10.3389/fpls.2023.1152036.

[37] Gandhi A, Kuknor S. The potential of virtual internships for management education: capturing student journey through experiential learning. Educ + Train 2024;66(2/3):341–54.

[38] Werner J, Jeske D. Ten simple rules for running and managing virtual internships. 2021,

[39] Zahariev P, Hristov G, Kinaneva D, Georgiev G. A pilot initiative for international virtual internships in the area of higher education. In: 2022 31st annual conference of the European association for education in electrical and information engineering. EAEEIE, IEEE; 2022, p. 1–4.

[40] Ezeafulukwe C, Bello BG, Ike CU, Onyekwelu SC, Onyekwelu NP, Asuzu OF. Inclusive internship models across industries: an analytical review. Int J Appl Res Soc Sci 2024;6(2):151–63.