Advancing Education through Tutoring Systems: A Systematic Literature Review

Vincent Liu, Ehsan Latif, and Xiaoming Zhai*

Abstract—This study systematically reviews the transformative role of Tutoring Systems, encompassing Intelligent Tutoring Systems (ITS) and Robot Tutoring Systems (RTS), in addressing global educational challenges through advanced technologies. As many students struggle with proficiency in core academic areas, Tutoring Systems emerge as promising solutions to bridge learning gaps by delivering personalized and adaptive instruction. ITS leverage artificial intelligence (AI) models, such as Bayesian Knowledge Tracing and Large Language Models, to provide precise cognitive support, while RTS enhance social and emotional engagement through human-like interactions. This systematic review, adhering to the PRISMA framework, analyzed 86 representative studies. We evaluated the pedagogical and technological advancements, engagement strategies, and ethical considerations surrounding these systems. Based on these parameters, Latent Class Analysis was conducted and identified three distinct categories: computer-based ITS, robot-based RTS, and multimodal systems integrating various interaction modes. The findings reveal significant advancements in AI techniques that enhance adaptability, engagement, and learning outcomes. However, challenges such as ethical concerns, scalability issues, and gaps in cognitive adaptability persist. The study highlights the complementary strengths of ITS and RTS, proposing integrated hybrid solutions to maximize educational benefits. Future research should focus on bridging gaps in scalability, addressing ethical considerations comprehensively, and advancing AI models to support diverse educational needs.

Index Terms—Tutoring Systems, Artificial Intelligence (AI), Intelligent Tutoring Systems (ITS), Robot Tutoring Systems (RTS), Personalized Learning, Engagement Strategies, Hybrid Educational Technologies

I. INTRODUCTION

MPROVING academic performance remains a global challenge, with many students struggling to achieve proficiency in core academic areas. For instance, a recent study highlighted that 96% of students in an urban Pennsylvania school performed below proficiency levels, while none reached proficiency at another site [1]. These findings underscore the critical need for innovative educational interventions that address systemic learning gaps.

Tutoring Systems, encompassing both *Intelligent Tutoring Systems (ITS)* and *Robot Tutoring Systems (RTS)*, are emerging as transformative educational technologies. ITS are software-based systems leveraging artificial intelligence (AI) to deliver adaptive, personalized instruction in real time [2], [3]. RTS extend these capabilities by integrating robotics to provide physical and social interactivity, fostering engagement through tangible interactions [4], [5]. Collectively, Tutoring Systems have demonstrated significant potential to enhance learning

The authors are with the AI4STEM Education Center, University of Georgia, Athens, GA, 30602, USA.

Corresponding Author Email: xiaoming.zhai@uga.edu

outcomes by adapting to individual student needs, promoting motivation, and addressing the limitations of traditional instructional methods [6], [7].

The significance of this study lies in its systematic evaluation of Tutoring Systems, particularly in the context of their impact on student engagement, academic achievement, and pedagogical practices. Research reveals that ITS can achieve learning gains comparable to one-on-one human tutoring [8], [9], while RTS have shown promise in improving student motivation and social-emotional learning outcomes [10], [11]. The integration of generative AI into ITS further enhances their adaptability, enabling real-time customization of instructional content based on student behavior and feedback [12], [13]. Similarly, RTS provide unique opportunities for collaborative learning and problem-solving, bridging the digital divide and fostering inclusivity in education [14], [15].

The evolution of tutoring systems, from rudimentary Computer Assisted Instruction (CAI) programs in the 1970s to today's advanced ITS and RTS, has transformed personalized learning. However, while technological breakthroughs have enabled unprecedented levels of adaptive instruction, critical questions remain regarding the practical impact on learning outcomes and the ethical implications of these innovations. Addressing these issues is essential for guiding future research and ensuring that educational technologies contribute positively to student development.

Despite their potential, Tutoring Systems face challenges, including ethical concerns related to AI bias, data privacy, and equitable access [16], [17]. This study seeks to address these issues by evaluating advancements in AI-driven tutoring, examining their adaptability to real-time inputs, and exploring best practices to ensure their ethical use [18], [19].

Below are the research questions address in this study:

- 1) How have tutoring systems evolved from the early CAI models of the 1970s to today's AI-driven ITS and RTS, and which technological breakthroughs have been most critical in enabling personalized and adaptive learning?
- 2) In what specific ways do modern tutoring systems incorporate real-time adaptive mechanisms, and how does this dynamic responsiveness translate into improved student engagement and academic performance across various educational contexts?
- 3) What ethical challenges, such as data privacy, informed consent, and algorithmic bias, have emerged alongside the technological advancements in tutoring systems, and what proven strategies have been developed to mitigate these risks while ensuring equitable and responsible use of these tools?

TABLE I
COMPARISON OF ITS AND RTS AS SUBCOMPONENTS OF TUTORING SYSTEMS

Dimension	Intelligent Tutoring Systems (ITS)	Robot Tutoring Systems (RTS)		
Pedagogical Focus	Cognitive and instructional personalization through	Social and emotional engagement via physical pres-		
	adaptive algorithms and AI models.	ence, gestures, and verbal communication.		
Adaptability	High adaptability with real-time adjustments to con-	Adaptability through interactive behaviors such as		
	tent, difficulty, and feedback using advanced AI	task adjustments and social responses, but primarily		
	techniques like Bayesian Knowledge Tracing (BKT)	pre-programmed and less data-driven.		
	and Large Language Models (LLMs).			
Personalization	Tailors instructional content using detailed student	Enhances interaction by recognizing students, using		
	models to address individual learning preferences	names, and offering social cues such as verbal praise		
	and weaknesses.	and gestures.		
Interaction	Text-based or conversational interfaces; cognitive	Multimodal interactions including physical gestures,		
	and information-focused interactions.	verbal communication, and eye contact, mimicking		
		human tutors.		
Engagement	Cognitive engagement driven by adaptive content	Emotional and social engagement through human-		
	and feedback, maintaining mental stimulation.	like behaviors, enhancing motivation and relational		
		connection.		
Scalability	High scalability due to reliance on existing computer	Limited scalability due to the need for physical		
	infrastructure and cloud-based deployment.	robotic hardware and associated maintenance.		
Ethical Considerations	Minimal ethical concerns; primary issues involve	Higher ethical concerns related to data privacy,		
	data privacy and ensuring fair AI-driven decision-	long-term human-robot interaction effects, and trans-		
	making.	parency in AI usage.		
Deployment Challenges	Easy deployment on existing digital platforms with	Complex deployment requiring significant invest-		
	minimal hardware requirements.	ment in robotic hardware and integration into phys-		
		ical environments.		

This paper is structured to systematically explore the role of Tutoring Systems in education. The introduction outlines the motivation and objectives of the study, followed by a detailed framework that characterizes Tutoring Systems. The methodology section describes the systematic review process, employing the PRISMA framework to identify and select relevant studies. The results section addresses the research questions, analyzing the impact of Tutoring Systems on educational outcomes, engagement, and ethical considerations. The discussion highlights key pedagogical and technological gaps, identifies integration challenges, and proposes future research directions. The paper concludes with a synthesis of findings and recommendations for advancing the field of Tutoring Systems.

II. INTELLIGENT TUTORING SYSTEM (ITS) VS. ROBOT TUTORING SYSTEMS (RTS)

This section provides a comparative analysis of Tutoring Systems based on their pedagogical effectiveness, adaptability, personalization, interaction methods, engagement strategies, and technical implementation. The comparison highlights their respective strengths and limitations, offering insights into how these systems address diverse educational needs and contexts.

1) Pedagogical Effectiveness: Tutoring Systems exhibit distinct pedagogical strengths. Software-based systems excel in delivering personalized and adaptive instruction by leveraging advanced AI models such as Bayesian Knowledge Tracing and Large Language Models (LLMs) [20]. These systems provide precise feedback and track progress to ensure cognitive engagement, which significantly enhances learning outcomes. However, they often lack the social and emotional engagement that contributes to holistic education.

Conversely, robot-based systems prioritize social and emotional engagement through physical presence, verbal encouragement, and human-like gestures [10], [21]. These features

are particularly effective for younger learners or those who benefit from interactive and relational learning. While robotbased systems support motivation and engagement, they face challenges in delivering deeply personalized and cognitively adaptive content at the same level as software-based systems.

2) Adaptability: Adaptability is a defining feature of software-based tutoring systems, achieved through dynamic adjustments to content and difficulty levels based on real-time student performance [22]. These systems utilize AI-driven techniques to provide highly individualized support, ensuring students are consistently challenged at appropriate levels.

Robot-based systems demonstrate adaptability primarily through interactive behaviors, such as providing hints or altering task complexity based on student responses [23]. However, this adaptability is often limited to pre-programmed responses, making it less flexible and nuanced compared to software-based systems. While robots excel in maintaining engagement, their cognitive adaptability remains an area for further development.

3) Personalization: Software-based systems personalize learning by analyzing detailed student profiles, including performance data and learning preferences, to tailor content and feedback effectively [20]. This approach ensures a targeted learning experience that addresses individual strengths and weaknesses.

Robot-based systems personalize interactions through social and emotional engagement, such as recognizing students by name and using verbal and non-verbal cues to foster a sense of connection [21]. However, their personalization is more focused on enhancing relational aspects rather than deeply customizing instructional content, limiting their cognitive personalization capabilities.

4) Interaction: The interaction styles of Tutoring Systems differ significantly. Software-based systems rely on text or conversational interfaces powered by AI models, offering efficient

and context-aware cognitive engagement [24]. However, their interactions remain screen-based, lacking the immersive and physical elements of human-like interaction.

Robot-based systems provide rich multimodal interactions, combining verbal communication, gestures, and eye contact to create a more personal and engaging learning experience [21], [23]. These physical and social interactions enhance motivation and emotional connection, making them particularly effective for younger learners.

5) Engagement: Engagement in software-based systems is primarily cognitive, driven by adaptive content and immediate feedback that keeps learners mentally stimulated [16]. While effective in maintaining focus and promoting academic achievement, these systems may lack the emotional resonance required for sustained motivation.

Robot-based systems excel in emotional and social engagement by mimicking human behaviors, offering encouragement, and fostering a sense of companionship [11]. This approach increases students' willingness to participate and enhances their attitudes toward learning, though it may not provide the same depth of cognitive engagement as software-based systems.

6) Technical Implementation: Software-based systems are relatively easy to deploy and scale due to their reliance on existing infrastructure and minimal hardware requirements. Cloud-based solutions further simplify updates and management, making these systems cost-effective and accessible in diverse educational settings.

Robot-based systems face greater deployment challenges, requiring significant investments in robotic hardware and supporting infrastructure. Scalability is limited by the cost and complexity of maintaining additional units. Hybrid approaches, such as combining robots with virtual avatars, offer potential solutions but may compromise the immersive experience robots are designed to provide [5], [15].

This comparative analysis underscores the complementary nature of software-based and robot-based Tutoring Systems. While software-based systems excel in cognitive adaptability and scalability, robot-based systems bring unique strengths in emotional and social engagement. Together, these systems can address diverse educational needs, offering opportunities for hybrid solutions that combine their respective advantages.

III. FRAMEWORK FOR ANALYZING TUTORING SYSTEMS

This review introduces a framework designed to analyze the evolution of tutoring systems across different generations. The framework is structured around four dimensions: Artificial Intelligence, system behavior, engagement, and affordances. To ensure a comprehensive analysis, representative papers from each generation of tutoring systems were carefully selected based on their alignment with these dimensions. These papers, identified through an extensive review of the literature, exemplify the defining characteristics and advancements of their respective generations. By systematically applying this framework to representative studies, the review provides a nuanced evaluation of the technological and pedagogical evolution of tutoring systems, highlighting their strengths, limitations, and potential for future integration (Fig. 1).

To systematically assess the progression of tutoring systems, we developed a framework that categorizes their evolution across different generations. This framework was constructed based on key technological advancements, pedagogical methodologies, and system capabilities observed in the literature. Therefore, we have structured the framework aroud four dimensions: Artificial Intelligence, system behavior, engagement, and affordances. As educational technologies have advanced, new capabilities, such as adaptive learning, AIdriven feedback, and robotic interaction, have emerged, fundamentally altering how students engage with tutoring systems. These shifts have been driven by advancements in computing power, machine learning, and human-computer interaction, yet there is limited work that systematically maps these developments into a structured framework. By introducing this generational model, we provide a historical context that allows for a clearer understanding of how tutoring systems have evolved to address different educational needs. Furthermore, analyzing tutoring systems through this generational lens strengthens the investigation into the impact of AI-driven advancements, their adaptability, and their ethical implications. Situating our findings within this broader trajectory enables a more comprehensive assessment of tutoring systems, ultimately offering insights that can inform future research and development in the field.

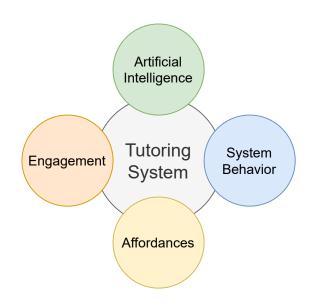


Fig. 1. Framework for Tutoring System Analysis

A. Artificial Intelligence

In the context of tutoring systems, Artificial Intelligence encompasses the algorithms and computational models that enable these technologies to deliver personalized and adaptive learning experiences [5], [13]. Early tutoring systems relied on foundational techniques, such as concept trees and student models, to map student progress and adapt instruction

based on predefined rules [25]–[27]. Subsequent advancements introduced Bayesian Knowledge Tracing, which improved predictive capabilities, allowing systems to adjust instructional content dynamically based on real-time student performance [22].

Modern tutoring systems integrate advanced AI models, such as ChatGPT and other large language models, to create interactive, conversational learning experiences [23], [24]. These models facilitate real-time feedback and adaptive interactions, tailoring the learning process to individual student needs. This evolution has transformed tutoring systems into responsive, AI-driven tools that emulate many qualities of human tutors, offering highly adaptive and personalized educational experiences [12].

B. System Behavior

System behavior refers to the adaptability of tutoring systems—their ability to modify instructional strategies and content in response to student inputs. Early systems offered limited adaptability, primarily delivering static content with minimal interaction. Over time, these systems evolved to include features such as adjusting problem difficulty, providing personalized feedback, and introducing tailored challenges based on student progress [21].

Advanced tutoring systems now leverage real-time analytics to dynamically adapt task complexity, provide hints, and deliver context-sensitive feedback [10], [23]. These capabilities enable systems to align more closely with students' learning paces and styles, fostering a personalized educational experience. By mirroring the flexibility of human tutors, modern systems enhance the learning process, supporting improved engagement and outcomes [5].

C. Engagement

Engagement in tutoring systems encompasses the strategies and features designed to maintain students' active involvement in the learning process [16]. Early systems relied on basic text-based feedback and limited interaction, offering minimal engagement [26], [27]. As these systems evolved, they incorporated more interactive elements, such as animated feedback [28], subgoal scaffolding [28], and real-time monitoring of concentration levels [29].

Contemporary tutoring systems utilize sophisticated engagement methods, including verbal communication, gestures, and physical movements that simulate human-like interactions [21], [23]. These systems can hold natural conversations, use facial expressions, and adapt dynamically to student responses, creating immersive and motivating learning environments. By integrating these advanced features, modern tutoring systems sustain student interest and motivation, making learning experiences more engaging and effective [10].

D. Affordances

Affordances in tutoring systems refer to the physical and interactive features that enhance student engagement and interaction. Early tutoring systems were primarily software-based,

providing limited physical interaction. Over time, systems began incorporating tangible elements, such as virtual avatars or robotic components, to create more immersive learning experiences [14].

Modern systems leverage physical features like robotic gestures, digital representations, and virtual environments to foster engagement and accessibility. For example, robotic tutors can use movement and physical presence to enhance interaction, particularly for younger learners or those who benefit from kinesthetic learning styles [11], [15]. These affordances bridge the gap between digital and physical learning, offering diverse and effective educational tools that cater to a wide range of learning needs [5].

IV. METHOD

This systematic review adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, proposed by Moher et al. [30], to ensure a transparent, replicable, and comprehensive approach to data collection and analysis. The PRISMA framework facilitated a structured process for identifying, screening, and selecting studies related to Tutoring Systems. The method comprises the following stages: identification, screening, eligibility, and inclusion, as depicted in Figure 2.

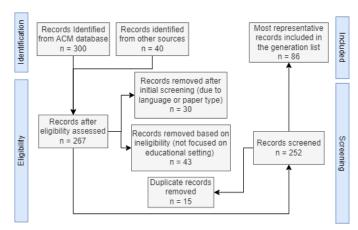


Fig. 2. Tutoring Systems PRISMA review flowchart

A. Search Strategy

To ensure transparency and replicability, this study adopted the PRISMA framework. As shown in the flowchart in Figure 2, we followed a systematic review process including overview of identification, screening, eligibility, and inclusion stages. To ensure a robust and comprehensive search, multiple bibliographic databases were used, including ACM, JSTOR, ResearchGate, BERA, ScienceDirect, Taylor & Francis, arXiv, ERIC, IOSPress, MDPI, EPFL, ISEDJ, and the American Psychological Association. Search queries focused on titles, abstracts, and keywords, using combinations of terms such as "Tutoring Systems," "adaptive learning," "AI in education," and "personalized learning."

Table II summarizes the search criteria applied to filter the studies. The initial search yielded a total of 340 records, ensuring inclusivity across article types, including research articles, journals, conference papers, and short papers.

To capture the full evolution of tutoring systems, this review spans a broad historical range from the early 1970s to the present. The earliest paper identified in our search, "CAI supports individualized learning programs in Montgomery County (Maryland) Public Schools" (1972), represents one of the pioneering works in computer-assisted instruction. By including studies from 1972 through 2024, we ensure that our analysis covers the inception of individualized learning technologies, the emergence and growth of ITS and RTS, and the latest innovations incorporating LLMs. This comprehensive time frame allows us to trace the technological, pedagogical, and ethical transformations that have shaped tutoring systems over more than five decades.

TABLE II SEARCH CRITERIA

Components	Examples	
Bibliography Databases	ACM, JSTOR, ResearchGate, BERA, Sci-	
	enceDirect, Taylor & Francis, arXiv, ERIC,	
	IOSPress, MDPI, EPFL, ISEDJ, American	
	Psychological Association	
Article Types	Research Articles, Short Papers, Journals,	
	Conference Papers	
Searched on	Title, Keywords, Abstract	
Sorting on Returns	Relevance and Time Period	
Language	English	

B. Inclusion and Exclusion Criteria

A systematic approach was used to establish inclusion and exclusion criteria, as summarized in Table III. Studies were included if they were in English, focused on educational settings, and involved experimentation or review of Tutoring Systems. Non-English documents, papers without educational relevance, or those without sufficient experimentation were excluded. Additional criteria ensured the inclusion of studies explicitly addressing Tutoring Systems with substantial educational relevance.

TABLE III
INCLUSION AND EXCLUSION CRITERIA FOR TUTORING SYSTEMS

Inclusion	Exclusion		
Documents in English	Documents not in English		
Focuses on educational settings	Not focused on educational settings		
Papers with experimentation or re-	Papers without experimentation		
views			
Studies addressing Tutoring	Studies with loose or no mention		
Systems with significant	of Tutoring Systems relevance		
teaching/learning focus			
Doc types: Research articles, jour-	Other doc types: Book reviews,		
nals, conference papers, short pa-	posters, etc.		
pers			

C. Study Selection Process

Identification: An initial search yielded 340 records, with 290 retrieved from the ACM Digital Library and 50 identified through additional searches on Google Scholar and Bing, following these procedures:

Screening: After duplicate removal, 252 unique records remained. This phase eliminated 15 duplicate records. Titles and abstracts were screened for relevance to Tutoring Systems in educational settings.

Eligibility: The eligibility evaluation removed 30 papers due to language barriers or document type incompatibilities (e.g., book reviews or editorials). An additional 43 records were excluded for lacking direct relevance to Tutoring Systems in educational contexts. Consequently, 179 records proceeded to the next stage.

Inclusion: After full-text assessment, 86 papers were included in the final review. These studies were deemed the most representative of their respective domains, aligned with the research objectives, and suited to the analytical framework.

While the review focuses on English-language publications, this criterion may have excluded relevant studies published in other languages. Additionally, the reliance on digital databases may introduce bias against non-digitized or regionally focused studies. This method ensures a thorough exploration of Tutoring Systems, providing a solid foundation for analyzing their impact on education.

D. Data Analysis

A systematic data extraction template was developed to according to our analytical framework. Key data points including but not limited to the study objectives, methodologies, sample sizes, educational settings, technologies employed, evaluation metrics, and findings, were coded by the authors. Interater reliability were monitored to make sure that the data coded were reliable. The extracted data were synthesized to identify trends, compare approaches, and highlight gaps in the current literature.

The methodological rigor ensures that this review captures a diverse range of perspectives on Tutoring Systems. By leveraging multiple databases and applying stringent inclusion criteria, the review provides a comprehensive overview of the field, highlighting both advancements and areas requiring further exploration.

E. Variables Measured

Ethics in tutoring systems is categorized into three levels: no ethical concerns, minimal ethical concerns, and high ethical concerns. Ethical concerns can range from data privacy to fears of abusing the AI to cheat. Papers classified as having no ethical concerns either explicitly state the absence of ethical issues or do not identify any potential risks associated with their tutoring system. For example, [23] does not mention any ethical concerns, and the system described does not present any apparent ethical risks. Papers with minimal ethical concerns acknowledge some potential issues but provide solutions or safeguards to address them. For instance, [31] highlights a feature allowing users to delete chat history, offering a degree of control over data privacy. In contrast, papers with significant ethical concerns identify more serious risks, such as overreliance on AI or its potential for generating misleading information. A clear example is [32], which warns readers

about the dangers of AI hallucinations potentially misleading students.

Adaptability refers to a system's ability to adjust instructional strategies in real time, linking closely to the dimensions of system behavior and AI. Papers are categorized as either having adaptive or fixed difficulty. Systems with adaptive difficulty, such as [33], dynamically modify teaching strategies and content based on student performance, creating a personalized learning experience. In contrast, systems with fixed difficulty, like [34], function more as one-way communication tools. These systems present linear content and difficulty levels without incorporating any input from the student, limiting their ability to respond to individual needs.

Interaction tools can significantly shape engagement and affordances in tutoring systems, spanning modalities such as computer software, tablet devices, and physical robots. Systems using computers as their primary interaction medium typically rely on software that acts as the student's tutor. For instance, in [20], students used software to learn mathematics, with the tutor dynamically personalizing content to meet individual student needs in real time. Similarly, systems employing tablet-like devices often integrate robots as complementary tools to enhance communication. For example, [23] describes a system where a tablet presents questions to students while a robot provides positive and encouraging feedback based on their answers. In contrast, systems that prioritize physical robots as the primary interaction tool offer more hands-on engagement. These robots often leverage gestures, movements, voices, and lifelike physical features to create immersive experiences. An example can be found in [21], where robots used a diverse range of gestures and movements to interact with students physically. The appearance of a tutoring system can also play a role in shaping the overall student experience. These systems can be categorized into non-anthropomorphic, semi-anthropomorphic, and fully anthropomorphic designs. Non-anthropomorphic systems are typically computers or tablet-like devices, which lack human-like features. Semianthropomorphic systems incorporate some human-like aspects but are missing key components. For example, [23] describes a robot with a human-like head and shoulders but without a full body. In contrast, fully anthropomorphic robots closely resemble human figures. An example is [35], which uses a NAO robot that resembles a human more closely, with features designed to be friendly and approachable.

Feedback mechanisms, such as text-based or verbal responses, play a crucial role in driving engagement and positively influencing a student's attitude toward the system. Feedback can be categorized into two main types: text-based and verbal. Text-based systems deliver feedback through written messages, typically displayed on a computer or tablet screen. For example, [36] describes a system that provides personalized feedback, including knowledge explanations, error analysis, and review suggestions, all conveyed through text on the student's screen. In contrast, verbal feedback systems use spoken responses, often delivered by a robot, to communicate with the student. These systems often incorporate encouraging language to motivate students and offer explanations through a different medium. An example of this is [23], where the robot

repeats the words displayed on the tablet, providing students with an additional layer of interaction and learning beyond just text on the screen.

AI techniques like Bayesian Knowledge Tracing (BKT) and Large Language Models (LLMs) are at the heart of these tutoring systems. BKT is a probabilistic model used in educational technology to track and predict a student's knowledge over time and can be seen in [37] where they use BKT to predict the student's knowledge level and adjust its content accordingly. LLMs are advanced AI systems trained on vast text data to understand and generate human language. In intelligent tutoring systems, LLMs provide personalized feedback, facilitate natural language interactions, generate adaptive content, and help students understand concepts through detailed explanations and dynamic dialogues. This can be seen in [38] where LLMs provide dynamic, personalized, and contentaware responses.

V. FINDINGS

A. Latent Class Analysis of Features of Tutoring Systems

Using categorical variables such as scalability, ethics, adaptability, device used, appearance, feedback, and AI techniques, we conducted Latent Class Analysis (LCA). Although the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) suggest that models with less latent classes provide better statistical fit, we believe that the 3-class model is the most appropriate for our scenario. This model strikes a balance between interpretability and capturing distinct characteristics of the latent classes. Adding more or reducing classes introduces complexity without providing substantial additional insight for this specific context.

The three primary classes are: Tutoring Systems relying on computer-based environments (ITS), Verbal Tutoring Systems using robot-based interactions, and Multimodal Tutoring Systems integrating multiple interaction modes. Each class exhibits distinct characteristics that align with specific educational affordances and pedagogical applications. This classification provides insights into the design choices and capabilities of tutoring systems and their impact on learning outcomes, engagement, and adaptability.By aligning these findings with the proposed framework, the analysis highlights how the dimensions of AI, system behavior, engagement, and affordances are shaping modern tutoring systems.

B. Categorical Features of Tutoring Systems

This section articulated the key features of each latent class. 1) Class 1: Computer-Based Tutoring Systems (ITS): Computer-based tutoring systems are characterized by non-anthropomorphic appearances and primarily text-based feedback. They rely heavily on AI techniques such as BKT and LLMs to deliver adaptive and personalized learning experiences [23], [25]. These systems demonstrate high adaptability, with a strong emphasis on real-time adjustments to content difficulty and feedback based on student performance. However, their scalability is moderate, and ethical concerns related to data privacy are generally less emphasized. An example of an ITS that strongly fits this category is [39],

where the system was entirely text-based, with students interacting solely through their computers. The system employed Bayesian Knowledge Tracing (BKT) to track student progress and dynamically adjust content to meet individual needs and difficulty levels. The computer interface was non-anthropomorphic, and the paper raised no ethical concerns.

- 2) Class 2: Verbal Tutoring Systems: Verbal tutoring systems are typically robot-based, with semi-anthropomorphic appearances and a focus on visual and verbal feedback. These systems enhance engagement by leveraging physical interactions and social behaviors, such as gestures and verbal cues, to create a more interactive learning environment [10], [21]. Although verbal systems demonstrate some adaptability, their reliance on pre-programmed content and external devices, such as tablets, limits their scalability and cognitive adaptability compared to computer-based systems. Ethical concerns are more prominent in this class, given the reliance on sensory data and potential risks associated with prolonged humanrobot interactions. An example of a system that exemplifies the verbal tutoring system category is [40], where the robot speaks and displays questions for the student to answer. The robot has limited mobility, with slight movements in its head, arm, and torso, making it semi-anthropomorphic. The system adapts dynamically by adjusting its questions in real time, using Large Language Models (LLMs) as its underlying AI technique. The paper also raises significant ethical concerns regarding LLM hallucinations, cautioning students about relying on them for factual information.
- 3) Class 3: Multimodal Tutoring Systems: Multimodal tutoring systems integrate advanced AI with highly scalable and interactive features. They employ anthropomorphic designs and verbal feedback, fostering a highly engaging and immersive learning experience [5], [15]. These systems utilize a variety of AI techniques, including LLMs and other advanced models, to adapt dynamically to student inputs. Multimodal systems excel in engagement and adaptability, offering a combination of verbal, visual, and physical interactions. However, their high complexity raises ethical concerns related to data privacy, transparency, and equitable access. An example of a system that highlights a multimodal tutoring system is [41] where the robot is fully anthropomorphic and equipped with a wide range of gestures and actions that make the NAO feel lifelike. The system uses BKT and can adapt its question in real time similar to other systems. The system did not have any ethical concerns and the feedback was a mix of verbal commands, gestures, and text from and external device.

This classification of tutoring systems demonstrates the diverse ways in which educational technologies can be tailored to meet specific learning needs. The latent class analysis reveals the strengths and limitations of each class, providing insights into how Tutoring Systems can address scalability, ethical considerations, adaptability, and engagement. By aligning these findings with the proposed framework, this analysis highlights areas for future development, particularly in enhancing cognitive adaptability and addressing ethical challenges.

4) Advancements in Artificial Intelligence: Tutoring Systems have witnessed substantial advancements in Artificial

Intelligence, enabling more personalized and adaptive learning experiences [5], [13]. Early systems relied on fundamental AI techniques such as concept trees and Bayesian Knowledge Tracing to track student progress and adjust content delivery [25]. Recent systems have adopted sophisticated models, including large language models (LLMs), to enhance responsiveness and interactivity [23]. These AI-driven models allow systems to dynamically tailor instructional content, deliver real-time feedback, and facilitate natural language interactions.

Modern Tutoring Systems also integrate Natural Language Processing (NLP) technologies, which support conversational tutoring and improve the overall student experience [10]. For example, LLM-based "tutor models" provide tailored instruction, while "feedback models" offer context-sensitive responses, significantly improving engagement and learning outcomes [9]. These advancements position Tutoring Systems as powerful tools for adaptive learning, with the capability to emulate many aspects of human tutoring.

5) System Behavior and Adaptability: The adaptability of Tutoring Systems is a cornerstone of their effectiveness. Alpowered algorithms such as Bayesian Knowledge Tracing and LLMs enable these systems to modify instructional strategies in real-time, responding to individual student needs [16]. For instance, modern systems dynamically adjust task difficulty, provide immediate feedback, and introduce new challenges based on a student's progress, ensuring an optimal balance of support and challenge [23].

The physical and social affordances of Tutoring Systems further enhance their adaptability. Robotic components within some systems enable real-time interactions, such as gestures, verbal praise, and physical task demonstrations, fostering engagement and motivation [10], [15]. However, cognitive adaptability—adjusting instructional content based on complex real-time analysis—remains more advanced in AI-driven software-based systems compared to robotic ones. While robots excel in maintaining engagement, their instructional capabilities often rely on external devices or pre-programmed content [5].

Studies comparing student performance with and without Tutoring Systems have demonstrated their potential to enhance learning outcomes and engagement. However, while software-based systems show significant improvements in academic achievement, the direct cognitive impact of robotic systems remains an area for further exploration.

6) Engagement through Interactive Features: Engagement is a critical dimension of Tutoring Systems, supported by features that actively involve students in the learning process [16]. Early systems provided basic interaction through static text and simple playback functions, which evolved into more dynamic tools incorporating animated feedback, verbal communication, and real-time monitoring of student attention [28], [29].

Modern systems integrate sophisticated methods, such as conversational AI and robot-driven gestures, to create immersive and interactive learning environments [23]. These features maintain student interest and motivation over extended periods, making the systems more effective than traditional static instructional methods [11], [21]. Advanced Tutoring Systems now employ subtle expressions, such as gazing or facial cues, which simulate human behavior, further enhancing

Dimension	Variable	Category	Computer-Based	Verbal	Multimodal
	Adaptability	Adaptive difficulty	40%	20%	40%
System Behavior		Fixed difficulty	50%	50%	30%
		Other	10%	30%	30%
	Interaction Tool	Computer	100%	0%	10%
Affordances		Tablet	0%	0%	60%
		Physical Robot	0%	100%	20%
	Appearance	Non-anthropomorphic	100%	0%	0%
		Semi-anthropomorphic	0%	40%	20%
		Full anthropomorphic	0%	40%	80%
Engagament	Feedback	Text-based	90%	50%	0%
Engagement		Verbal	10%	50%	100%
	AI Technique	BKT	20%	0%	40%
		LLM	70%	60%	40%
Artificial Intelligence		Other	10%	40%	20%
	Responsible AI	Ethical concerns	20%	40%	40%
		Minimal concerns	20%	20%	0%

No concerns

TABLE IV
LATENT CLASS ANALYSIS: CHARACTERISTICS OF TUTORING SYSTEM CLASSES

TABLE V
COMPARISON OF AIC AND BIC FOR DIFFERENT NUMBERS OF LATENT
CLASSES

Number of Classes (k)	AIC	BIC
1	-134.54	-128.57
2	-786.13	-773.19
3	-1075.37	-1055.45
4	-1168.59	-1141.71
5	-1255.41	-1221.55

the quality of engagement and aligning with diverse student needs.

C. Ethical Considerations and Responsible AI Use

The rapid adoption of AI-powered Tutoring Systems raises critical ethical concerns, particularly regarding data privacy, security, and fairness [17]. Despite collecting extensive student data, many studies lack a thorough examination of how this sensitive information is managed and protected. This highlights a significant gap in the literature and underscores the need for robust ethical frameworks.

To address these challenges, key measures for responsible AI use include implementing data encryption protocols, conducting regular security audits, and establishing stringent access controls to safeguard student data [42]. Transparency is also essential, with educational institutions and developers required to clearly communicate data collection and usage practices to students and guardians. Mitigating AI biases through diverse training datasets and fairness-aware algorithms is equally critical. Regular audits and human oversight are necessary to ensure equitable treatment and maintain accountability.

Moreover, governance frameworks with oversight committees can support the ethical deployment of Tutoring Systems, ensuring compliance with established data protection standards and promoting trust in AI-driven education [9]. By integrating these measures, Tutoring Systems can foster a secure, equitable, and effective learning environment.

This review highlights the transformative potential of Tutoring Systems, driven by advancements in Artificial Intelligence,

enhanced adaptability, and interactive engagement features. While these systems offer promising solutions to educational challenges, ethical considerations remain a crucial area for development. Future research should focus on bridging gaps in robotic cognitive adaptability, addressing ethical concerns comprehensively, and exploring long-term impacts on student learning outcomes.

40%

60%

60%

VI. DISCUSSION

This study highlights the transformative potential of Tutoring Systems in addressing educational challenges through advanced Artificial Intelligence, adaptability, and interactive engagement. Software-based systems demonstrate exceptional cognitive adaptability and scalability, leveraging AI models such as Bayesian Knowledge Tracing and Large Language Models (LLMs) to deliver personalized, adaptive learning experiences [20], [23]. These systems are particularly effective in tracking student progress, tailoring content, and providing real-time feedback, which supports robust cognitive engagement and improved learning outcomes.

Robot-based systems, on the other hand, excel in fostering emotional and social engagement through human-like behaviors such as gestures, verbal encouragement, and physical interactions [10], [21]. These systems are highly effective in maintaining student motivation and creating immersive learning environments, especially for younger learners or those requiring more relational approaches [5]. Despite these strengths, both systems exhibit distinct limitations, such as the lack of holistic engagement in software-based systems and the scalability challenges faced by robot-based systems.

The integration of software- and robot-based Tutoring Systems offers significant potential to address these limitations. By combining the cognitive adaptability of software-based systems with the social and emotional engagement of robot-based systems, integrated solutions could provide a more comprehensive educational experience. Such systems could dynamically adapt content and interactions to both cognitive performance and emotional states, promoting deeper learning and sustained motivation [15].

Despite these promising findings, several gaps remain that limit the full potential of Tutoring Systems. Future research should address these gaps.

Pedagogical Gaps: While software-based systems excel in cognitive engagement, their reliance on text-based or screen interactions fails to accommodate diverse learning styles, particularly kinesthetic and auditory learners [16]. Robot-based systems, although effective in fostering engagement, often lack the cognitive adaptability and precision required for delivering highly personalized instructional content [23]. Both systems need to adopt a more integrated pedagogical approach that addresses the cognitive, emotional, and experiential dimensions of learning.

Technological Gaps: Data privacy and security remain critical concerns, with many Tutoring Systems collecting extensive student data but lacking robust protocols for its protection [16], [42]. The integration of LLMs, while enhancing adaptability, introduces risks such as potential misuse for academic dishonesty, necessitating improved monitoring mechanisms [9]. In robot-based systems, the reliance on dual-system designs (robots as adjuncts to tablets) highlights a lack of fully integrated AI capabilities, limiting their potential as autonomous educational tools [10].

Scalability and Accessibility Gaps: Software-based systems, while scalable, face challenges in reaching underresourced regions due to disparities in technological infrastructure and socio-economic barriers [5]. Robot-based systems encounter even more pronounced scalability challenges due to the high costs and logistical complexities of deploying robotic hardware in diverse educational settings [15]. Addressing these gaps requires the development of cost-effective solutions that ensure equitable access to advanced educational tools.

Ethical Considerations: Both types of systems need to prioritize ethical considerations, including data transparency, fairness, and accountability. The long-term social implications of human-robot interaction in education remain underexplored, raising concerns about dependency, equity, and the potential impact on student development [10], [21].

Addressing these gaps will require a multifaceted approach: **Holistic Integration:** Developing systems that integrate the cognitive strengths of software-based systems with the social and emotional capabilities of robot-based systems can create balanced learning environments tailored to diverse needs.

Enhanced AI Models: Improving the capabilities of AI to operate seamlessly within robot-based systems will enhance their adaptability and instructional capabilities, moving beyond adjunctive designs [23].

Scalable Solutions: Designing cost-effective, accessible systems for deployment in under-resourced settings can reduce disparities in educational technology access.

Ethical Frameworks: Establishing transparent guidelines for data collection, usage, and security will build trust and ensure equitable treatment of all students.

By addressing these challenges, future Tutoring Systems can become more effective, equitable, and transformative, significantly contributing to the advancement of global education.

VII. CONCLUSION

The evolution of tutoring systems highlights their transformative potential in advancing educational outcomes through personalized and adaptive learning experiences. Softwarebased systems, exemplified by ITS, excel in cognitive adaptability, leveraging advanced AI techniques to deliver tailored instruction. Conversely, RTS brings unique strengths in engagement and motivation through physical presence and human-like interactions, making them particularly effective for relational and immersive learning. This review suggests that integrating ITS and RTS offers a promising path forward. By combining the cognitive precision and scalability of ITS with the interactive and engaging qualities of RTS, hybrid systems could address a broader range of student needs. Such systems have the potential to dynamically adapt to both academic performance and social-emotional cues, providing tailored support and feedback in real time. This integration could enhance learning outcomes, particularly for students who benefit from multi-sensory and experiential learning environments.

ACKNOWLEDGMENT

This study secondary analyzed data from projects supported by the Institute of Education Sciences (Grant Number R305C240010, PI Zhai). The authors acknowledge the funding agencies and the project teams for making the data available for analysis. The findings, conclusions, or opinions herein represent the views of the authors and do not necessarily represent the views of personnel affiliated with the funding agencies.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work the author(s) used ChatGPT in order to check grammar and polish the wordings. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

REFERENCES

- [1] D. R. Thomas, J. Lin, E. Gatz, A. Gurung, S. Gupta, K. Norberg, S. E. Fancsali, V. Aleven, L. Branstetter, E. Brunskill *et al.*, "Improving student learning with hybrid human-ai tutoring: A three-study quasiexperimental investigation," in *Proceedings of the 14th Learning Analytics and Knowledge Conference*, 2024, pp. 404–415.
- [2] E. Mousavinasab, N. Zarifsanaiey, S. R. Niakan Kalhori, M. Rakhshan, L. Keikha, and M. Ghazi Saeedi, "Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods," *Interactive Learning Environments*, vol. 29, no. 1, pp. 142–163, 2021.
- [3] J. A. Kulik and J. D. Fletcher, "Effectiveness of intelligent tutoring systems: a meta-analytic review," *Review of educational research*, vol. 86, no. 1, pp. 42–78, 2016.
- [4] S.-T. Chu, G.-J. Hwang, and Y.-F. Tu, "Artificial intelligence-based robots in education: A systematic review of selected ssci publications," *Computers and education: Artificial intelligence*, vol. 3, p. 100091, 2022.
- [5] X. Zhang, F. Zhu, K. Wang, G. Cao, Y. Xue, and M. Liu, "Bring the intelligent tutoring robots to education: A systematic literature review," *IEEE Transactions on Learning Technologies*, 2024.
- [6] W. Ma, O. O. Adesope, J. C. Nesbit, and Q. Liu, "Intelligent tutoring systems and learning outcomes: A meta-analysis." *Journal of educa*tional psychology, vol. 106, no. 4, p. 901, 2014.

- [7] S. Feng, A. J. Magana, and D. Kao, "A systematic review of literature on the effectiveness of intelligent tutoring systems in stem," in 2021 IEEE frontiers in education conference (fie). IEEE, 2021, pp. 1–9.
- [8] K. Holstein, B. M. McLaren, and V. Aleven, "Intelligent tutors as teachers' aides: exploring teacher needs for real-time analytics in blended classrooms," in *Proceedings of the seventh international learning analytics & knowledge conference*, 2017, pp. 257–266.
- [9] S. Maity and A. Deroy, "Generative ai and its impact on personalized intelligent tutoring systems," arXiv preprint arXiv:2410.10650, 2024.
- [10] A. Ramachandran, C.-M. Huang, and B. Scassellati, "Toward effective robot-child tutoring: Internal motivation, behavioral intervention, and learning outcomes," ACM Transactions on Interactive Intelligent Systems (TiiS), vol. 9, no. 1, pp. 1–23, 2019.
- [11] K.-Y. Chin, Z.-W. Hong, and Y.-L. Chen, "Impact of using an educational robot-based learning system on students' motivation in elementary education," *IEEE Transactions on learning technologies*, vol. 7, no. 4, pp. 333–345, 2014.
- [12] S. Banjade, H. Patel, S. Pokhrel et al., "Empowering education by developing and evaluating generative ai-powered tutoring system for enhanced student learning," Journal of Artificial Intelligence and Capsule Networks, vol. 6, no. 3, pp. 278–298, 2024.
- [13] C.-C. Lin, A. Y. Huang, and O. H. Lu, "Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review," Smart Learning Environments, vol. 10, no. 1, p. 41, 2023.
- [14] J. Yang and B. Zhang, "Artificial intelligence in intelligent tutoring robots: A systematic review and design guidelines," *Applied Sciences*, vol. 9, no. 10, p. 2078, 2019.
- [15] K. Wang, G.-Y. Sang, L.-Z. Huang, S.-H. Li, and J.-W. Guo, "The effectiveness of educational robots in improving learning outcomes: A meta-analysis," *Sustainability*, vol. 15, no. 5, p. 4637, 2023.
- [16] K. Swargiary, "The impact of ai-driven personalized learning and intelligent tutoring systems on student engagement and academic achievement: Ethical implications and the digital divide," Available at SSRN 4897241, 2024
- [17] A. Robert, K. Potter, and L. Frank, "The impact of artificial intelligence on students' learning experience," *Logical Techniques in computer* science, vol. 2, no. 2, pp. 71–76, 2024.
- [18] Y. Akyuz et al., "Effects of intelligent tutoring systems (its) on personalized learning (pl)," Creative Education, vol. 11, no. 06, p. 953, 2020
- [19] T. Ahmadov, A. Karimov, S. Durst, M. Saarela, W. Gerstlberger, M. F. Wahl, and T. Karkkainen, "A two-phase systematic literature review on the use of serious games for sustainable environmental education," *Interactive Learning Environments*, pp. 1–22, 2024.
- [20] S. E. Fancsali and S. Ritter, "Context personalization, preferences, and performance in an intelligent tutoring system for middle school mathematics," in *Proceedings of the fourth international conference on learning analytics and knowledge*, 2014, pp. 73–77.
- [21] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro, "Interactive robots as social partners and peer tutors for children: A field trial," *Human–Computer Interaction*, vol. 19, no. 1-2, pp. 61–84, 2004.
- [22] Š.-G. Ines, G. Ani, and G. Angelina, "Twenty-five years of bayesian knowledge tracing: a systematic review," *User Modeling and User-Adapted Interaction*, pp. 1–47, 2024.
- [23] E. Verhelst, R. Janssens, T. Demeester, and T. Belpaeme, "Adaptive second language tutoring using generative ai and a social robot," in Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, 2024, pp. 1080–1084.
- [24] A. Lieb and T. Goel, "Student interaction with newtbot: An Ilm-as-tutor chatbot for secondary physics education," in *Extended Abstracts of the* CHI Conference on Human Factors in Computing Systems, 2024, pp. 1–8.
- [25] V. J. Shute and J. Psotka, "19. intelligent tutoring systems: Past, present, and future," 1994.
- [26] G. Hume, J. Michael, A. Rovick, and M. Evens, "The use of hints by human and computer tutors: the consequences of the tutoring protocol," 1996.
- [27] T. W. Draper and W. W. Clayton, "Using a personal robot to teach young children," *The Journal of genetic psychology*, vol. 153, no. 3, pp. 269–273, 1992.
- [28] A. Corbett and H. Trask, "Instructional interventions in computer-based tutoring: Differential impact on learning time and accuracy," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2000, pp. 97–104.
- [29] G. Lee, A. Ojha, and M. Lee, "Concentration monitoring for intelligent tutoring system based on pupil and eye-blink," in *Proceedings of the*

- 3rd International Conference on Human-Agent Interaction, 2015, pp. 291–294.
- [30] D. Moher, L. Shamseer, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, L. A. Stewart, and P.-P. Group, "Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015 statement," *Systematic reviews*, vol. 4, pp. 1–9, 2015.
- [31] C. Dong, "How to build an ai tutor that can adapt to any course and provide accurate answers using large language model and retrievalaugmented generation," arXiv preprint arXiv:2311.17696, 2023.
- [32] H. Kumar, I. Musabirov, M. Reza, J. Shi, X. Wang, J. J. Williams, A. Kuzminykh, and M. Liut, "Impact of guidance and interaction strategies for llm use on learner performance and perception," arXiv preprint arXiv:2310.13712, 2023.
- [33] C. L. King, Vincent, Kelvin, H. L. Warnars, N. Nordin, and W. H. Utomo, "Intelligent tutoring system: Learning math for 6th-grade primary school students," *Education Research International*, vol. 2021, no. 1, p. 5590470, 2021.
- [34] J. Kanero, E. T. Tunalı, C. Oranç, T. Göksun, and A. C. Küntay, "When even a robot tutor zooms: A study of embodiment, attitudes, and impressions," *Frontiers in Robotics and AI*, vol. 8, p. 679893, 2021.
- [35] A. Lekova, P. Tsvetkova, T. Tanev, P. Mitrouchev, and S. Kostova, "Making humanoid robots teaching assistants by using natural language processing (nlp) cloud-based services," *Journal of Mechatronics and Artificial Intelligence in Engineering*, vol. 3, no. 1, pp. 30–39, 2022.
- [36] Z. Jiang and M. Jiang, "Beyond answers: Large language model-powered tutoring system in physics education for deep learning and precise understanding," arXiv preprint arXiv:2406.10934, 2024.
- [37] Z. A. Pardos, M. Tang, I. Anastasopoulos, S. K. Sheel, and E. Zhang, "Oatutor: An open-source adaptive tutoring system and curated content library for learning sciences research," in *Proceedings of the 2023 chi* conference on human factors in computing systems, 2023, pp. 1–17.
- [38] B. Alsafari, E. Atwell, A. Walker, and M. Callaghan, "Towards effective teaching assistants: From intent-based chatbots to Ilm-powered teaching assistants," *Natural Language Processing Journal*, vol. 8, p. 100101, 2024.
- [39] M. Eryılmaz and A. Adabashi, "Development of an intelligent tutoring system using bayesian networks and fuzzy logic for a higher student academic performance," *Applied Sciences*, vol. 10, no. 19, p. 6638, 2020.
- [40] M. Helal, P. Holthaus, L. Wood, V. Velmurugan, G. Lakatos, S. Moros, and F. Amirabdollahian, "When the robotic maths tutor is wrong-can children identify mistakes generated by chatgpt?" in 2024 5th International Conference on Artificial Intelligence, Robotics and Control (AIRC). IEEE, 2024, pp. 83–90.
- [41] T. Schodde, K. Bergmann, and S. Kopp, "Adaptive robot language tutoring based on bayesian knowledge tracing and predictive decisionmaking," in *Proceedings of the 2017 ACM/IEEE international confer*ence on human-robot interaction, 2017, pp. 128–136.
- [42] C. A. Eden, O. N. Chisom, I. S. Adeniyi et al., "Integrating ai in education: Opportunities, challenges, and ethical considerations," Magna Scientia Advanced Research and Reviews, vol. 10, no. 2, pp. 006–013, 2024.