Acta Informatica Pragensia
2025, Volume 14, Issue 2, 215–245
Special Issue: Theory and Practice of Generative Artificial Intelligence Usage
https://doi.org/10.18267/j.aip.269

PRAGUE UNIVERSITY OF ECONOMICS AND BUSINESS

Article Open Access

Generative Artificial Intelligence in Ubiquitous Learning: Evaluating a Chatbot-based Recommendation Engine for Personalized and Context-aware Education

Manel Guettala [1], **Samir Bourekkache** [1], **Okba Kazar** [2], **Saad Harous** [3]

[1] Laboratoire de l'INFormatique Intelligente (LINFI), Department of Computer Science, University of Mohamed Khider Biskra, Algeria
[2] Department of Computer Science, College of Arts, Sciences, IT & Communication, University of Kalba, Sharjah, United Arab Emirates
[3] Department of Computer Science, College of Computing and Informatics, University of Sharjah, Sharjah, United Arab Emirates

Corresponding author: Manel Guettala (manel.guettala@univ-biskra.dz)

Editorial Record

First submission received:
October 9, 2024

Revisions received:
November 9, 2024
January 31, 2025
March 16, 2025
May 4, 2025
May 29, 2025

Accepted for publication:
May 30, 2025

Special Issue Editors:
David Chudan
Prague University of Economics and Business, Czech Republic

Miroslav Vacura
Prague University of Economics and Business, Czech Republic

This article was accepted for publication by the Special Issue Editors upon evaluation of the reviewers' comments.

How to cite this article:
Guettala, M., Bourekkache, S., Kazar, O., & Harous, S. (2025). Generative Artificial Intelligence in Ubiquitous Learning: Evaluating a Chatbot-based Recommendation Engine for Personalized and Context-aware Education. *Acta Informatica Pragensia*, 14(2), 215–245. https://doi.org/10.18267/j.aip.269

Abstract

Background: Ubiquitous learning environments aim to provide personalized and context-aware educational resources; however, traditional recommendation systems often fall short in meeting these dynamic learner needs.

Objective: This study develops and evaluates a chatbot-based recommendation system that uses generative AI and prompt engineering techniques to enhance recommendation accuracy and user engagement in ubiquitous learning contexts.

Methods: A ChatGPT-powered chatbot was implemented using few-shot prompting and dynamic context integration to deliver personalized, real-time educational support. The system was deployed using an intuitive Gradio interface, facilitating user accessibility and seamless interaction across varied learning scenarios. A tailored evaluation dataset was constructed to capture diverse user interactions and the system was tested through real-world case studies and user feedback metrics, including task success rates, response times and satisfaction ratings.

Results: The chatbot achieved an 85% overall task success rate, a 70% success rate in context-aware tasks and an 80% user satisfaction rating, with most users assigning scores of 4 or 5 on a 5-point scale.

Conclusion: The findings demonstrate that the proposed solution outperforms traditional systems in delivering personalized, adaptive and context-aware educational recommendations, underscoring the transformative potential of generative AI in advancing learner-centred ubiquitous learning environments.

Index Terms

Ubiquitous computing; Context awareness; Deep learning; Personalized educational; Educational technologies; HCI; Recommendation system; LLMs; Generative AI; ChatGPT; Prompting engineering; Few-shot prompting.

1 INTRODUCTION

Ubiquitous computing facilitates the availability of educational resources at any time and in any location, thereby necessitating the development of recommendation systems that are customized to specific contexts (Guettala et al., 2021; Guettala et al., 2022). Context-aware recommender systems (CARS) fulfil this requirement by adaptively modifying recommendations based on contextual elements such as temporal and spatial variables (Adomavicius and Tuzhilin et al., 2010).

These systems employ sophisticated methodologies, including pre-filtering, post-filtering and contextual modelling, to amalgamate user, item and contextual data to deliver personalized recommendations (Haruna et al., 2017). As articulated by Remountakis et al. (2023), recommendation technologies (RT) harness the capabilities of artificial intelligence and machine learning to forecast user inclinations and mitigate the challenges associated with information overload. Techniques such as supervised learning (k-nearest neighbours (Peterson et al., 2009), decision trees (Kingsford et al., 2008) and unsupervised learning (k-means clustering (Likas et al., 2003) facilitate clustering of users or classification of data points. CARS enhances these methodologies by integrating contextual elements, thus providing dynamic and customized recommendations.

The objective of ubiquitous learning environments is to provide personalized and context-aware educational resources, aligning with the overarching goal of enhancing learning outcomes (Guettala et al., 2024a). This study focuses on learning as the central theme, exploring how large language models, such as ChatGPT, can address limitations in traditional recommendation systems by delivering adaptive and context-sensitive suggestions.

We do this by answering the following research questions:

- **RQ1:** To what extent do chatbot-driven recommendations enhance the accuracy, relevance and user satisfaction of personalized learning in ubiquitous learning environments?
- **RQ2:** In what ways did integrating natural language interaction in our chatbot affect user engagement and perceived effectiveness in ubiquitous learning environments compared to traditional recommendation systems?
- **RQ3:** How effectively did the chatbot adapt to dynamic user contexts (e.g., location, time and preferences) and optimize response generation for greater accuracy and relevance in ubiquitous learning environments?

This study explores the integration of our chatbot, powered by ChatGPT, within ubiquitous learning environments, offering novel insights into the evolving role of AI in education. Central to this exploration are two key concepts: personalization, which refers to tailoring recommendations based on individual user preferences, learning history and goals, and context sensitivity, which involves adapting recommendations based on situational factors such as location, time and user preferences. These concepts work in tandem to ensure that the chatbot delivers highly relevant and adaptive recommendations, addressing both users' long-term learning objectives and their immediate situational needs.

The paper is organized as follows: Section 2 reviews the literature, analysing prior research and debating existing AI-driven recommendation systems in education. Section 3 outlines the methodology and system design, presenting the development of our chatbot as a recommendation system tailored for adaptive, context-aware learning experience. Section 4 examines real-world case studies, demonstrating the effectiveness of our chatbot compared to traditional recommendation systems. Section 5 focuses on data collection methods, utilizing key metrics. The evaluation is integrated with existing literature to contextualize the findings and highlight our chatbot's contributions. Finally, this section answers the research questions, synthesizing the findings. Section 6 concludes the study by summarizing its contributions and proposing future research directions to further refine our chatbot and expand its role in advancing personalized education.

2 LITERATURE REVIEW

Seamless availability of educational resources is a fundamental aspect of ubiquitous learning, and recommendation systems play a vital role in personalizing and optimizing these resources to meet individual learners' needs. This section highlights the various recommendation algorithms and systems, specifically those designed for ubiquitous learning, and connects them to the broader objectives of this study. It also explicitly identifies existing research gaps and clarifies how the present study advances beyond prior work.

2.1 Classification of recommendation systems

Recommendation systems are categorized based on their underlying technological principles and functional features. The most common approaches include:

- **Content-based filtering:** This approach recommends items by analysing the features of previously selected content (Lops et al., 2011). Frameworks such as TF-IDF and cosine similarity (Pazzani and Billsus, 2007) are commonly used. While effective in personalization, it suffers from cold start problems when new users or items are introduced (Adomavicius and Tuzhilin, 2005).

- **Collaborative filtering:** This method predicts user preferences based on similarities between users or items (Schafer et al., 2007). Matrix factorization techniques, such as singular value decomposition (Koren et al., 2009) and *k*-nearest neighbours (Sarwar et al., 2001), are widely adopted but can struggle with scalability and data sparsity (Koren et al., 2021).
- **Hybrid models:** These systems combine content-based and collaborative filtering to overcome their individual limitations (Burke, 2002). The Netflix recommendation algorithm (Gomez-Uribe and Hunt, 2015) is a well-known example. However, hybrid models often require complex computational resources (Fernández et al., 2012).
- **Context-aware systems:** These systems enhance recommendations by incorporating contextual factors such as time, location and user preferences (Adomavicius and Tuzhilin, 2005). Context-aware recommendation models, such as the mobile recommender system (Baltrunas et al., 2012), improve personalization but face challenges related to data availability and real-time adaptability (Verbert et al., 2012).

2.2 Context-aware recommendation systems

El Guabassi et al. (2016) presented a context-sensitive recommendation framework that tailors educational resources utilizing decision tree methodologies. By employing *k*-fold cross-validation, they validated the system effectiveness in delivering accurate recommendations that correspond with the unique contexts of learners. However, this approach relied on fixed decision pathways, limiting adaptability to the evolving and fluctuating needs of learners.

Expanding on this foundation, Thiprak and Kurutach (2015) developed a context-aware system for Thai herbal studies, using QR code scanning to deliver customized recommendations. This multimodal approach enhanced educational outcomes but lacked dynamic adaptability beyond its predefined dataset.

Similarly, Wang and Wu (2011) applied association rule mining to lifelong learning, yet their framework remained constrained by static learning pathways rather than incorporating real-time learning adaptation.

Durán and Álvarez (2017) developed a smartphone-based recommendation system for students in computer networks, which allows real-time access to learning resources via QR codes. However, its limited scalability reinforced the need for AI-driven solutions that dynamically adjust recommendations based on ongoing interactions rather than predefined rules.

2.3 Generative AI in recommendation systems

Recent advancements in generative AI models such as ChatGPT have significantly improved the adaptability and personalization of recommendation systems. These models, built on the transformer architecture (Waswani et al., 2017), make use of self-attention mechanisms to efficiently process sequential data, making them well-suited for dynamic and interactive learning applications.

In Figure 1, a series of concentric circles illustrates the model hierarchy: at the centre is "ChatGPT," within the "GPT Model" layer, which sits inside the "LLM Model" layer, all encompassed by the broader "Generative AI" domain. This diagram highlights ChatGPT as a specific instance of GPT models, which use the transformer architecture, enabling powerful language understanding and generation. Positioned within the generative AI field, ChatGPT showcases the transformative potential of transformers for creating adaptive and personalized learning systems (Guettala et al., 2023).

Guettala et al. (2024b) demonstrated the ability of ChatGPT to provide real-time, human-like responses, showcasing its potential in delivering personalized and context-aware learning experience. Similarly, Iatrellis et al. (2024) explored its integration into academic advising, demonstrating its effectiveness in enhancing decision-making with personalized recommendations.

Sondakh et al. (2024) developed a web-based recommendation system using GPT-3 to assist students in selecting research topics, highlighting the role of personalization in improving educational outcomes. Deldjoo et al. (2024) explored generative models in recommender systems (Gen-RecSys), emphasizing engagement and adaptability through AI-generated content. These studies reinforce the potential of generative AI in education but do not fully explore the role of real-time adaptability based on contextual factors, which is central to this study.

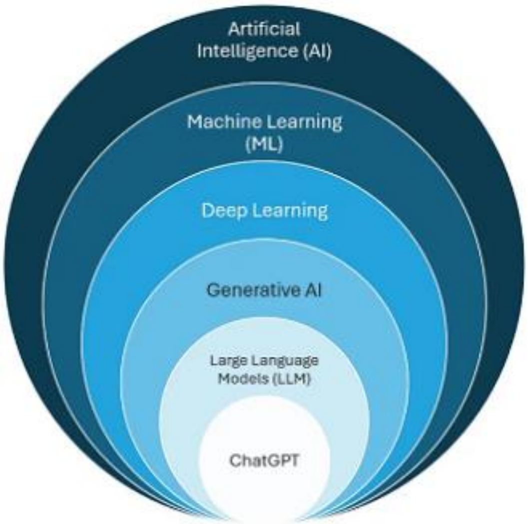

Figure 1. *Positioning of ChatGPT within hierarchy of generative AI models.*

2.4 User satisfaction in recommendation systems

User satisfaction is a critical metric for evaluating recommendation system effectiveness, particularly in conversational AI applications. The ability of a system to meet user expectations through relevance, responsiveness and adaptability directly influences its adoption.

Siro et al. (2023) examined conversational recommender systems, emphasizing dialogue relevance and efficiency as key satisfaction drivers. Zhang et al. (2024) explored user satisfaction in ChatGPT-based recommendation systems, finding that prompt guidance and domain-specific optimization significantly enhance user experience. Arapakis et al. (2021) investigated the role of response latency, showing that delays exceeding 7–10 seconds lead to frustration, underscoring the importance of real-time response in AI-driven systems.

2.5 Research gap and contribution of this study

While prior research has demonstrated the effectiveness of generative AI in personalized recommendations, significant gaps remain in real-time adaptability, dynamic learning adjustments and contextual responsiveness. Studies such as Sondakh et al. (2024) focus on static, rule-based recommendation models that fail to dynamically adjust to user interactions. Similarly, Deldjoo et al. (2024) explore chatbot usability in education but do not emphasize real-time learning adaptation based on evolving user preferences.

Recommendation system approaches, including content-based filtering (Lops et al., 2011), collaborative filtering (Pazzani and Billsus, 2007) and hybrid models (Burke, 2002), have significantly contributed to personalization in learning environments. However, these approaches often face challenges such as cold start issues, data sparsity and the inability to dynamically adapt to changing user contexts. Context-aware systems (Baltrunas et al., 2012) have attempted to address these limitations but struggle with real-time adaptability and reliance on static contextual inputs.

The present study advances beyond these limitations by integrating real-time NLP-based adaptability, allowing our chatbot to personalize recommendations dynamically according to user context (e.g., location, time constraints and learning preferences). Unlike previous works that relied on fixed learning pathways, our framework incorporates adaptive learning mechanisms that evolve through continuous user interactions, ensuring a more flexible, engaging and personalized learning experience.

Moreover, while prior research acknowledged the importance of personalized learning, adaptability was often treated as a secondary feature rather than a core component. In contrast, our study explicitly investigates how chatbot-driven recommendations dynamically adjust to user needs, differentiating it from earlier studies that primarily focused on predefined learning pathways.

Thus, this research presents a novel approach that combines generative AI with real-time learning adaptability, offering a scalable, interactive and context-aware recommendation system. By addressing the limitations of prior works, our study contributes to the advancement of AI-driven learning platforms with enhanced adaptability, personalization and real-time responsiveness.

3 METHODOLOGY

This study presents the design and implementation of a chatbot recommendation system underpinned by the GPT-3.5-turbo model, specifically tailored for ubiquitous learning environments. Unlike conventional recommendation engines that rely on rigid algorithmic frameworks or task-specific models, our system utilizes the inherent generalization capacity of large language models (LLMs) via prompt engineering. The chatbot autonomously adapts to a variety of learning scenarios and user profiles without requiring domain-specific fine-tuning. Core to this adaptability is the integration of few-shot prompting, which allows the system to incorporate diverse user contexts and deliver highly personalized and context-sensitive recommendations based on minimal training exemplars.

3.1 Global architecture of a GPT-driven chatbot for ubiquitous learning

The developed chatbot functions as a standalone recommendation system, interfacing with the GPT-3.5-turbo API via OpenAI's infrastructure. Rather than fine-tuning the underlying model, we employed a dynamic prompt construction strategy that uses structured templates and contextual embeddings to elicit tailored outputs. This design ensures platform independence, scalability and computational efficiency, critical for deployment across diverse ubiquitous learning settings. As illustrated in Figure 2, the system architecture comprises several key stages, including context acquisition, prompt assembly, language model inference and post-processing. This modular design enables flexible deployment and facilitates rigorous evaluation of its performance in varied educational use cases.

3.1.1 Task-specific prompt construction

To ensure that each recommendation is both personalized and pedagogically relevant, we developed an automated prompt construction module composed of three sequential components:

1. **User context integration:** Upon receiving a user query, the system captures and integrates contextual variables such as learning objectives, domain expertise, location, preferred content modality (e.g., text, video) and historical interactions. This information is either explicitly provided by the user or retrieved from session-level metadata, allowing the system to enrich the user input without requiring repeated manual input.
2. **Few-shot prompt embedding:** A curated repository of task-specific exemplars (few-shot examples) is maintained, each labelled according to educational domain and interaction type (e.g., tutoring, resource recommendation, career advice). When a query is submitted, the middleware dynamically selects representative exemplars aligned with the query’s intent and injects them into the system prompt. This strategy enables the chatbot to generalize from prior successful cases, facilitating zero-shot adaptation to novel tasks with minimal reliance on annotated data.
3. **Output format specification:** To standardize the format and improve the interpretability of the chatbot’s recommendations, a structured output template is appended to each prompt. This template instructs the model to generate responses with the following components:
   - *Title:* A succinct name for the recommended resource or action.
   - *Description:* A summary highlighting key features or learning objectives.
   - *Contextual justification:* An explanation linking the recommendation to the user's specific context (e.g., prior knowledge, goals or constraints).
   - *Actionable steps:* Clear guidance or direct links to access the recommended resource.

This structured prompt enables consistency across responses, enhances usability and facilitates post-hoc validation. Table 1 illustrates examples of prompt-response pairs across diverse educational tasks, demonstrating how the system customizes its output in accordance with user context and interaction type.

The methodological novelty lies in the adaptive orchestration of few-shot prompts based on user metadata, which enables scalable and precise personalization. This approach addresses key challenges in recommendation systems such as cold start problems, static behaviour and data sparsity. Moreover, it situates the system as a generative alternative to hybrid recommendation architectures by unifying personalization, context awareness and conversational fluency within a single LLM-based framework.

3.1.2 Chatbot-based recommendation generation and dialogue handling

Once the prompt is assembled, the system queries the GPT-3.5-turbo model to generate recommendations in real time. The process involves the following steps:

1. **Interactive query handling:** Users initiate natural language requests via a web-based interface. The system parses the query and, if necessary, requests additional clarification to ensure precise context capture.
2. **Contextual inference and natural language understanding:** Making use of the pretrained capabilities of GPT, the chatbot interprets the user's intent, preferences and learning objectives. The model uses embedded few-shot examples and context metadata to simulate a human-like understanding of the query.
3. **Personalized recommendation generation:** Based on the composite prompt, the model produces a structured response. Recommendations are dynamically adjusted in real time, with the system refining outputs through a feedback mechanism that incorporates user ratings, interaction history and the output refinement module (described in Section 3.2).

This dialogic interaction model positions the chatbot not merely as a static recommender but as an adaptive learning assistant capable of maintaining coherent and goal-aligned conversations across multiple turns.

3.1.3 Refinement and verification

The refinement and verification module is a component custom-developed by the authors, designed to ensure the accuracy, relevance, completeness and contextual appropriateness of outputs generated by the GPT-3.5-turbo model. Rather than modifying the base language model, this module operates as an independent post-processing layer that systematically evaluates and refines generated content to meet established quality standards.

a. Verification and refinement workflow

The module follows a structured, multi-stage workflow that includes:

1. **Output validation:** Each response is evaluated based on criteria defined by the authors, informed by instructional design principles and domain-specific requirements. The key validation checks include:
   - *Formatting checks:* Ensuring that outputs conform to a consistent, structured format (e.g., title, description, justification, actionable steps), validated against a predefined schema.
   - *Completeness checks:* Verifying that all required fields and content elements are present and fully populated.
   - *Relevance checks:* Assessing contextual alignment by comparing generated content with user-specific information, such as learning history, stated preferences and environmental factors (e.g., time or location).
   - *Factual accuracy verification:* Cross-referencing generated statements against curated domain-specific knowledge bases and external authoritative sources. Responses with a confidence score below 0.75 are flagged for manual review.
2. **Iterative corrections:** If any validation criteria are unmet, the system initiates an automatic reprocessing loop. Parameters are adjusted, such as prompt structure or example selection, and the query is resubmitted to the model. This cycle continues until the output passes all quality checks or reaches a predefined iteration limit, after which it is escalated for manual intervention.

3. **Finalization and learning adaptation:** Once validated, outputs are stored as reference exemplars for future interactions. This creates a growing repository of high-quality responses, which are used to enhance future outputs via few-shot prompting, contributing to the chatbot's adaptive capabilities.

b. Application and practical implementation

The functionality of the refinement and verification module is demonstrated in Section 4, where real-world case studies illustrate its application in healthcare education and professional development:

- In Section 4.1.1, the module refines diabetes management articles for a 45-year-old user by validating relevance, cross-referencing medical guidelines and ensuring readability before presenting them.
- In Section 4.1.2, the module performs personalization of continuing education resources for healthcare professionals based on variables such as medical specialty, years of experience and preferred learning modality, ensuring compliance with professional development standards.

c. Enhancing adaptive learning

This module plays a critical role in delivering personalized, high-quality educational content within ubiquitous learning environments. By enforcing rigorous output validation and using stored exemplars for continuous adaptation, the system ensures that recommendations are actionable, context-aware and pedagogically sound. The integration of few-shot prompting techniques further improves personalization and scalability, positioning the chatbot as an effective tool for adaptive, learner-centred instruction.

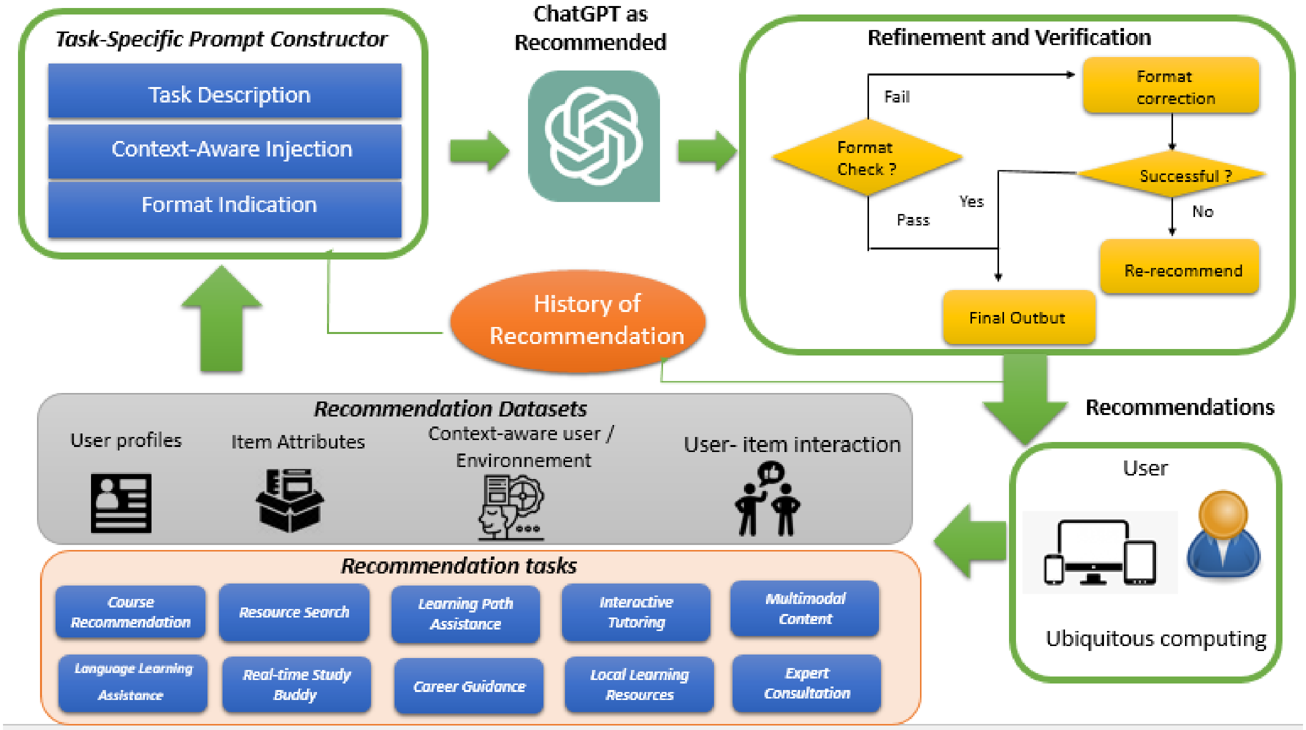

Figure 2. *Our process for ChatGPT-based recommender system for ubiquitous learning.*

3.2 Implementing a GPT-driven chatbot for ubiquitous learning

Building on the foundational methodology introduced in Section 3.1, this section explores the advanced design and functionality of our chatbot, which demonstrates significant potential to transform ubiquitous learning environments. The chatbot is powered by the GPT-3.5-turbo model (Abramski et al., 2023) with the OpenAI library, chosen for its remarkable ability to deliver context-aware and personalized recommendations. As depicted in Figure 4, this model addresses several limitations of its predecessors, such as those in the Instruct GPT series, offering superior performance and adaptability.

A key strength of GPT-3.5-turbo is its excellence in few-shot learning, which allows the chatbot to quickly adapt to new tasks with minimal training data. This capability ensures versatility and relevance across a wide range of recommendation scenarios. Additionally, the dynamic responsiveness to user queries and preferences enables the system to cater to the unique requirements of individual learners, significantly enhancing the overall user experience. Figure 3 provides a segment of the code implemented in our chatbot recommendation system.

```
def UbiquitousLearningChatbot(user_input):
    messages.append({"role": "user", "content": user_input})
    response = openai.ChatCompletion.create(
        model="gpt-3.5-turbo",
        messages=messages
    )
    chatbot_reply = response["choices"][0]["message"]["content"]
    messages.append({"role": "assistant", "content": chatbot_reply})
    return chatbot_reply
```

Figure 3. *Part of source code for our chatbot.*

The following step-by-step breakdown illustrates how the code functions to process user prompts and generate personalized recommendations.

a. Input processing

The function *UbiquitousLearningChatbot* begins by capturing the user's input, which is passed as the parameter *user_input*. The input is then appended to the *messages* list with the role of "user".

b. Prompt engineering

The *messages* list serves as the input prompt for the GPT-3.5-turbo model. The system dynamically constructs the context for the AI model based on this list, enabling the chatbot to generate a personalized response.

c. Model query and response

The *response* object contains the model output, which is then processed to extract the relevant response text. This step ensures that the chatbot retrieves the generated reply from the API result.

d. Output formatting

The chatbot appends the generated reply to the *messages* list with the role of "assistant". This ensures that the response is logged and can be used to maintain context for future interactions.

e. Refinement and verification

The function returns the chatbot's reply for display or further processing. This step ensures that the recommendations are clear, actionable and aligned with the user's needs.

This code is integral to the chatbot's ability to deliver real-time, context-aware recommendations. With dynamically processing user input and generating personalized responses, the chatbot enhances the learning experience in ubiquitous environments. The use of few-shot prompting and contextual data integration further improves the system scalability and adaptability, making it a powerful tool for personalized education.

Our solution is distinguished by its unparalleled accessibility, designed to be inclusive and available to anyone, anytime and anywhere. Aligned with the principles of ubiquitous learning, it ensures seamless usability across devices, including PCs, mobile phones and tablets. This is made possible through the Gradio interface (Abid et al., 2019), which simplifies access to machine learning models by enabling users to interact with the system via a shared URL.

The application depicted in Figure 4 is a chatbot-based recommendation system powered by GPT-3.5-turbo, designed to deliver personalized and context-aware educational resources in ubiquitous learning environments. The system is integrated with the Gradio interface, enabling users to interact with the chatbot in real time through a web-based platform. In Section 4.1, we provide a detailed description of the application features and use cases, showcasing its functionality and effectiveness in delivering personalized and context-aware recommendations. This section demonstrates how the system dynamically adapts to user needs, offering tailored learning experience that enhance engagement and outcomes in diverse educational contexts.

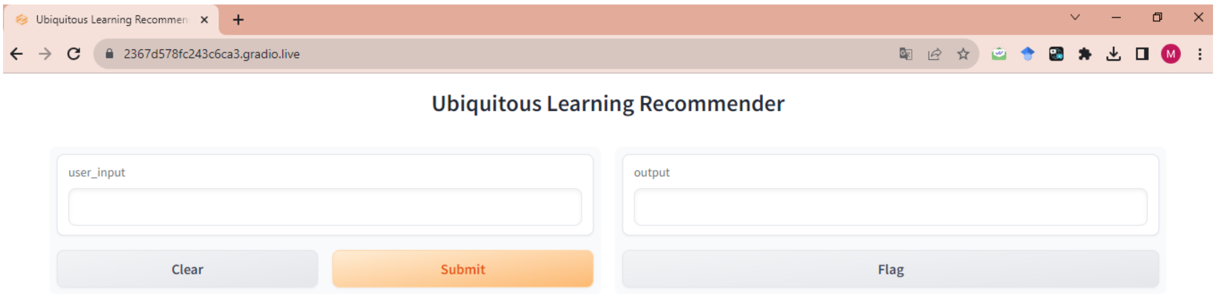

Figure 4. *Example of output.*

The prompts in Table 1 represent a subset of task-specific examples designed to demonstrate the chatbot's ability to deliver personalized and context-aware recommendations. While they do not cover all possible prompts, the system is highly adaptable and scalable, capable of handling diverse user requests through few-shot prompting and contextual data integration.

Table 1. *Summary of each task-specific inquiry.*

Task-specific prompt	**Example request**
Prompt for course recommendation (learning objects)	*"I'd like to broaden my knowledge of artificial intelligence. Could you recommend a good online course in this field for me?"*
Prompt for resource search	*"I'm looking for research papers on alternative energy sources. Could you assist me in locating some current publications?"*
Prompt for learning path assistance	*"I'm a high school student who wants to work in biology. Could you recommend a learning path that covers the fundamental concepts and resources that I should begin with?"*
Prompt for interactive tutoring	*"I'm having trouble understanding calculus. Can you provide me with step-by-step explanations and practice problems to help me improve?"*
Prompt for multimodal content recommendation (learning style)	*"I prefer video and interactive simulations for learning. Can you recommend any multimedia resources to help me understand the fundamentals of quantum physics?"*
Prompt for language learning assistance	*"I'm attempting to learn Spanish. Could you recommend a regular language practice schedule as well as any useful applications or websites?"*
Prompt for real-time study buddy (cooperative learning)	*"I have a math exam tomorrow and would like to study with someone online. Can you put me in touch with a study companion who is also studying the same subject?"*
Prompt for career guidance (skills)	*"I'm thinking about changing careers and going into data science. Could you please advise me on the required skills, certifications and job opportunities in this field?"*
Prompt for local learning resources (location context)	*"I live in New York City and would like to attend live coding workshops. Could you please suggest any upcoming events or coding meetups in my area?"*
Prompt for an expert consultation	*"I need help with my chemistry project. Can you recommend a local chemistry expert available for face-to-face consultation?"*

4 EXPERIMENTAL ANALYSIS AND DISCUSSION

In this section, we examine real-life case studies demonstrating how our chatbot, powered by ChatGPT, can be utilized as a recommendation system in healthcare education and other ubiquitous learning scenarios. The study uncovers various situations where the chatbot proves flexible and effective in meeting diverse learning goals by analysing extensive user interactions. By comparing our chatbot-based recommendation system with traditional approaches, this analysis highlights both its strengths and areas for improvement. The chatbot excels in delivering personalized, context-aware recommendations, adapting dynamically to user needs. However, certain limitations, such as handling resource-intensive or location-specific tasks and occasional response delays, present opportunities for further refinement. This evaluation underscores what works well in the system, such as dynamic interaction and context sensitivity, while identifying enhancements needed to maximize its potential in educational recommendation systems.

4.1 Case study approach

We present a real-world example in the healthcare education domain to illustrate how to use our ubiquitous learning-based recommendation system.

4.1.1 Scenario 1: Personalized health education recommendations

Our chatbot supports tailored guidance in ubiquitous learning by allowing users to refine their preferences through multi-turn, user-driven interactions. In this predominantly user-centric approach, users initiate questions and the chatbot provides customized recommendations that respond to individual needs. Using the methodology outlined in Section 3.1, a specific scenario illustrates how the chatbot adapts its guidance to meet personalized learning requirements.

Step 1: Task-specific prompt construction

The chatbot is utilized by a health education platform to enhance its recommendations for individuals seeking health information and guidance. It generates task-specific prompts tailored for health-related recommendations, following the three key components outlined in Section 3.1.1.

User information integration

The prompt includes essential details such as the user's health status, age and preferred content genre (e.g., articles or videos). For example: *"Good day, ChatGPT! You help people locate customized health education content. Please produce suggestions depending on the user's health status, age and preferred content genre (for example, articles or videos). The suggestions must follow current health guidelines and be medically accurate."*

Context-aware processing using few-shot responses

Few-shot prompting techniques are employed to simulate diverse user-item interactions. For instance, the prompt may include previous interactions where users with similar health conditions and preferences received tailored recommendations. This helps the chatbot infer patterns and generate contextually appropriate suggestions.

Output format specification

The chatbot's recommendations are structured to ensure clarity and actionability. The output format includes:

- **Title:** The name of the recommended resource (e.g., "Managing Diabetes: A Comprehensive Guide").
- **Description:** A brief summary of the recommendation, highlighting its key features or benefits (e.g., "This article provides practical tips for managing diabetes through diet and exercise.").
- **Contextual justification:** An explanation of why the recommendation is relevant to the user's context (e.g., "This article is tailored for individuals over 40 who are managing diabetes.").
- **Actionable steps:** Direct links or instructions for accessing the recommended resource (e.g., "Click here to read the article.").

Our chatbot enables users to specify their needs incrementally, facilitating more personalized recommendations. For example, as shown in Figure 5, users can provide additional preferences for tailored suggestions with context awareness, such as their health status (learning objective), age and preferred learning style (content genre).

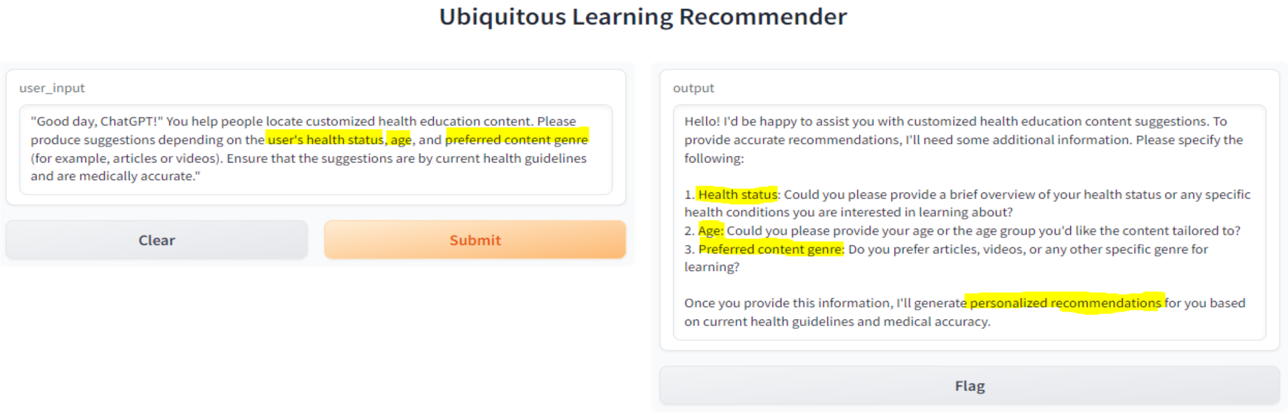

***Figure 5**. Prompt construction.*

Step 2: Process ChatGPT

Nazim logs into his account and seeks assistance from the chatbot. He shares the following personal details:

- Medical condition: Diabetes.
- Age: 45
- Preferred content type: Articles.

The chatbot initiates a real-time interaction with Nazim, allowing him to specify his preferences and context. The chatbot analyses Nazim's input, understands his request and provides a set of initial recommendations for diabetic health education. It suggests a curated list of engaging articles focused on various aspects of diabetes management, tailored specifically for individuals over the age of 40.

Step 3: Refinement and verification

The system runs the obtained findings via the refinement module, which conducts the following activities before giving the suggestions to Nazim:

- The module verifies the relevance of the user's age (45) to the health condition (diabetes) and ensures that the recommended articles are appropriate for the age range.
- The chatbot cross-references the recommendations with the most recent health guidelines and diabetes management practices to ensure medical accuracy and relevance.
- The module checks the readability and comprehension of the articles to ensure that they are suitable for Nazim's learning preferences. Once all quality criteria are met, the recommendations are finalized and presented to Nazim.

Consequently, in Figure 6, Nazim receives a list of high-quality articles that are tailored precisely to his health.

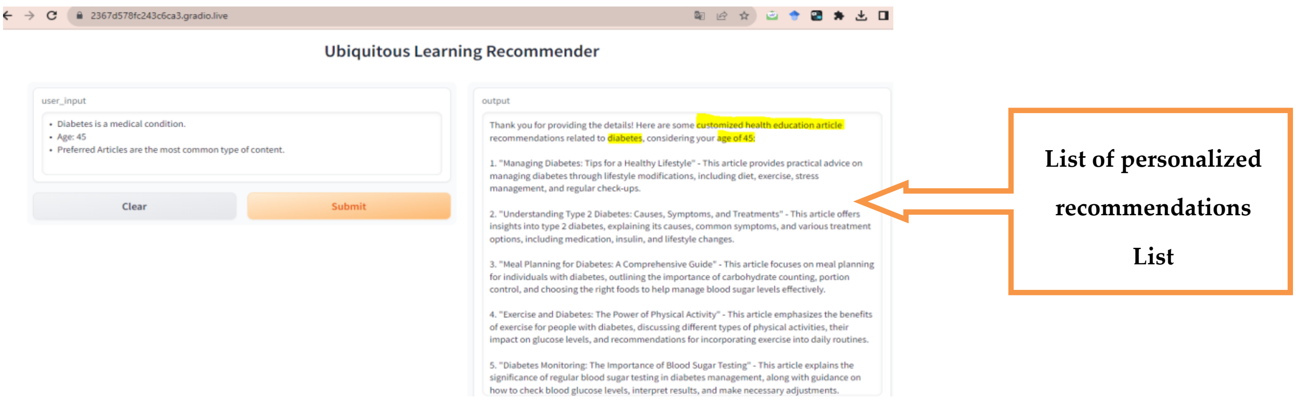

***Figure 6**. Set of preliminary recommendations for diabetic health education resources.*

4.1.2 Scenario 2: Ubiquitous learning for healthcare professionals

This scenario demonstrates how the chatbot supports healthcare professionals in accessing tailored resources for continuous education. Using the methodology outlined in Section 3.1, the chatbot generates personalized recommendations based on the user's medical specialty, years of experience and preferred learning format.

Step 1: Task-specific prompt construction

The chatbot is integrated into a healthcare training platform to provide personalized recommendations for continuing education. It generates task-specific prompts tailored for healthcare professionals, following the three key components outlined in Section 3.1.1.

User information integration

The prompt includes essential details such as the user's medical specialty, years of experience and preferred format (e.g., webinars, case studies). For example: *"Good day, ChatGPT! You help healthcare professionals locate suitable continuing education materials. Please make recommendations depending on the user's medical specialty, years of experience and preferred format (for example, webinars and case studies). Make certain that the recommendations are in line with current medical practices and professional development requirements."*

Context-aware processing using few-shot responses

Few-shot prompting techniques are employed to simulate diverse user-item interactions. For instance, the prompt may include previous interactions where healthcare professionals with similar specialties and experience levels received tailored recommendations. This helps the chatbot infer patterns and generate contextually appropriate suggestions.

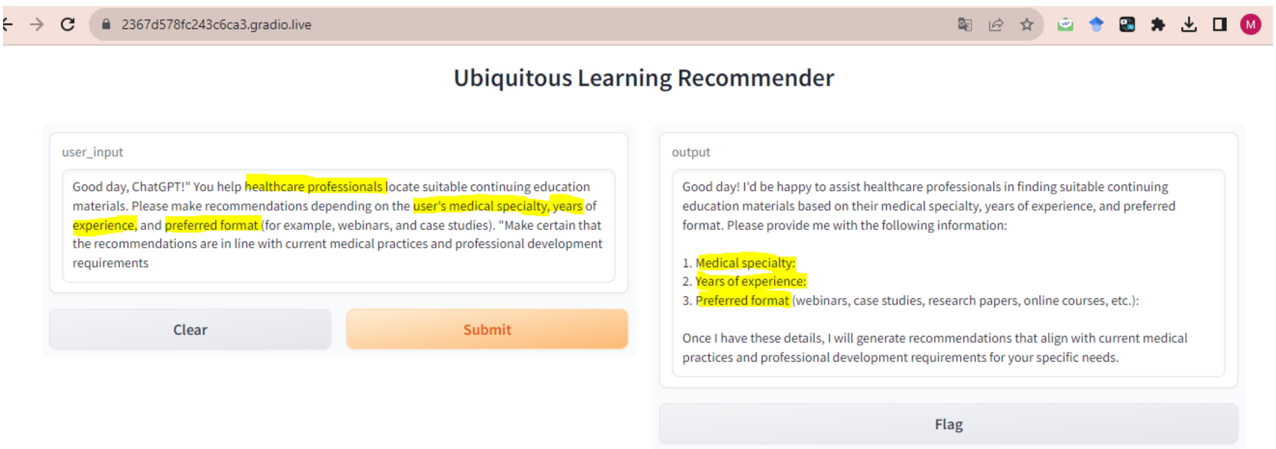

Figure 7. *Prompt construction.*

Output format specification

The chatbot's recommendations are structured to ensure clarity and actionability. The output format includes a title, description, contextual justification and actionable steps.

Step 2: Process ChatGPT

In Figure 7, Dr. Taha, a healthcare practitioner, seeks assistance from the chatbot. Dr. Taha provides the following details:

- Medical specialty: Cardiology.
- Experience: 12 years.
- Preferred format: Case studies.

The chatbot processes Dr. Taha's input, understands the request and delivers a tailored set of recommendations for continuing education in cardiology. It suggests a collection of advanced cardiology case studies designed to align with the expertise of a practitioner with 12 years of experience.

Step 3: Refinement and verification

The system runs the obtained findings via the refinement module, which performs the following activities before generating the suggestions to Dr. Taha:

- Validates the suitability of cardiology as a medical specialty for someone with 12 years of experience.
- Ensures that the necessary case studies are obtained from trustworthy institutions and are up to date with the most recent advances in cardiology.
- Checks the recommendations against current medical standards and industry best practices.
- Assesses the complexity of the case studies and depth of Dr. Taha's expertise.

Following these verification stages, the refining module validates the applicability of the recommendations. Dr. Taha is given the modified suggestions, which have now been proven accurate, relevant and aligned with universal learning goals for healthcare professionals.

Dr. Taha is presented with a selection of high-quality case studies that are suited to his medical specialty, degree of expertise and preferred format. This tailored approach benefits his professional growth by ensuring that the recommendations are not only relevant but also consistent with current medical practices, as shown in Figure 8.

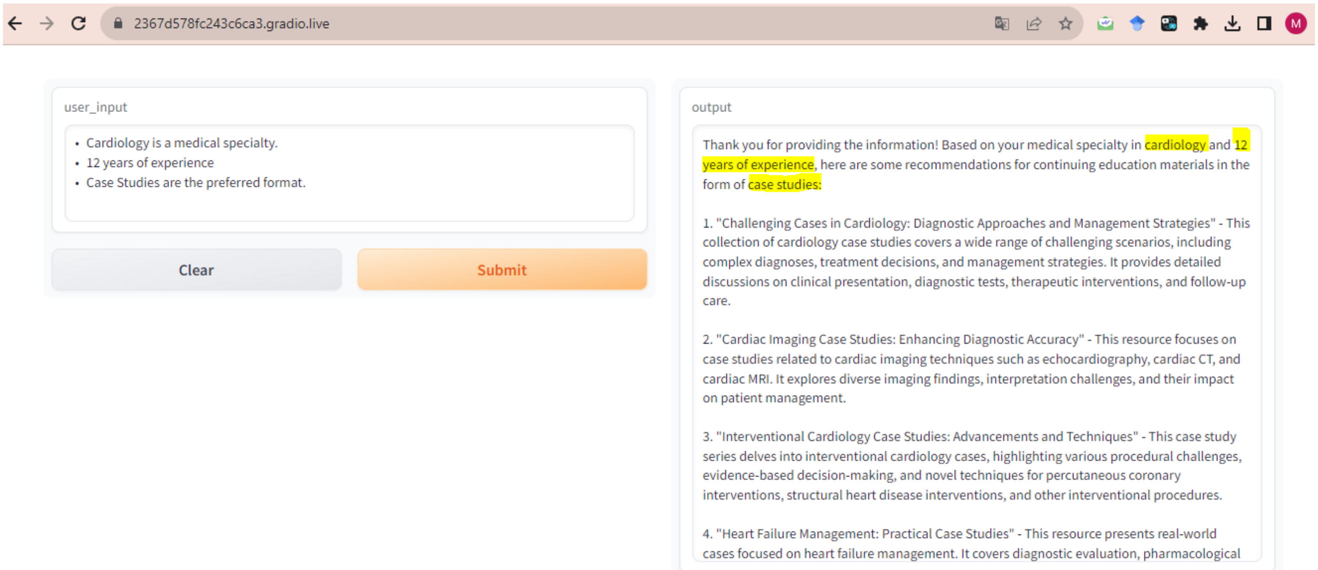

Figure 8. *List of personalized recommendations*.

Our chatbot recommendations are refined in this ubiquitous learning scenario for healthcare professionals to meet the critical criteria of medical accuracy and alignment with the user's specialization, experience and professional development needs, resulting in a valuable and effective resource for continuing education.

4.2 Discussion and analysis of personalization through our chatbot-driven recommendations

Section 4.1 highlights how personalization and context-aware suggestions can significantly enhance learning by adapting to individual needs and learning styles. By working with unique user data, our chatbot delivers more relevant recommendations, making learning more effective. This is particularly important in fields such as healthcare, where accuracy and applicability are essential. To ensure the correctness and relevance of suggestions, recommendations are refined through a verification process. This section discusses and analyses the impact of chatbot-based recommendation systems in ubiquitous learning environments.

Scenario 1: Personalized health education recommendations

In this scenario, our chatbot provided Nazim with recommendations for contextually appropriate health education. To provide relevant recommendations, the algorithm considered parameters such as the user's age (45), preferred content genre (articles) and health status (diabetes).

Scenario 2: Ubiquitous learning for healthcare professionals

Our chatbot made recommendations to Dr. Taha for context-aware continuing education in the context of healthcare practitioners. The algorithm tailored recommendations based on Dr. Taha's years of experience (12), the preferred format (case studies) and medical specialty (cardiology).

Our chatbot-driven recommendation systems have the potential to significantly enhance learning in ubiquitous environments by personalizing the learning experience. This system can improve the learning experience in several ways, as discussed in Table 2.

Table 2. *Personalization aspects of chatbot-driven recommendation systems in ubiquitous learning environments*.

Personalization aspect	**Description**
Personalization	Recommendations are tailored using user-specific data, ensuring relevance to individual interests, needs and learning goals.
Context awareness	Recommendations consider context factors such as health condition, age, specialization and experience level to align with the user's specific situation.
Refinement module	A refinement process ensures recommendations are accurate, reliable and adhere to professional standards and norms.
Adaptability	ChatGPT adapts to various fields (e.g., health education and professional development), offering tailored recommendations in different learning environments.

5 EVALUATION AND FINDINGS

This study examined the effectiveness of advanced language models, particularly our chatbot, as a recommendation system for ubiquitous learning. The results suggest that our chatbot holds the potential to significantly enhance the development and functionality of future educational material recommendation systems. Participants from 100 diverse backgrounds interacted with the chatbot, allowing us to evaluate its utility, user satisfaction and contextual awareness. The following sections detail the methods, metrics, dataset and insights derived from the user feedback and interactions.

5.1 Method

To conduct the evaluation, we utilized Gradio, a versatile UI library, to design a user-friendly interface that enabled participants to interact authentically with our ChatGPT-powered chatbot. Users accessed the chatbot through a seamless URL interface, which allowed them to engage in various learning tasks such as interactive tutoring, resource searching and course recommendations. Every interaction was meticulously recorded, capturing critical data such as task type, response time, context awareness, user satisfaction and task outcomes. This systematic data collection provided comprehensive insights into the chatbot's performance and the user experience across diverse academic and learning contexts.

5.2 Overview of dataset

The dataset consists of interactions from 100 individuals with diverse educational and professional backgrounds who engaged with our chatbot over a specified period. The dataset includes the six attributes outlined in Table 3, which were carefully selected to balance data collection efficiency with the ability to evaluate the chatbot's performance effectively. These attributes capture critical metrics that are central to assessing the effectiveness of the system in ubiquitous learning environments. The dataset is intentionally concise to ensure user privacy and streamline analysis.

Table 3. *Description of attributes in system evaluation dataset.*

Attribute	**Description**
User_ID	Each user has a unique identification.
Task_Type	User request type, such as course recommendation, resource search, interactive tutoring and others

Attribute	Description
Response_Time	Time taken by the chatbot (in seconds) to react to the user's request
Context_Aware	A binary value indicating whether the interaction was context-aware (“yes” or “no”)
User_Rating	User’s evaluation of the chatbot's response, rated on a scale from 1 (least satisfied) to 5 (most satisfied)
Task_Result	The outcome of the interaction, classified as "success" or "failure"

Before responding, the Context_Aware attribute denotes whether the chatbot considered any contextual data, such as the user's profile, location or learning history. If the chatbot’s response was influenced by contextual information, the value for this attribute is "yes". If the response was based solely on the user’s query without any additional context, the value is recorded as "no"”.

The Task_Result attribute was determined through a combination of user feedback and system evaluation. If the chatbot provided a recommendation and the user confirmed its relevance or completed the suggested task, the outcome was marked as "success”. If the user rejected the response, found it irrelevant or did not act on it, the outcome was recorded as "failure”. In addition, the refinement module played a role in validating task success by cross-checking the relevance and contextual accuracy of the chatbot-generated responses before presenting them to the user.

5.3 Data collection method

To ensure comprehensive data collection, we implemented a systematic approach to track and log user interactions with our ChatGPT-based recommendation system. The following is a comprehensive description of the data collection procedure, accompanied by illustrative code snippets for recording user interactions.

5.3.1 User interaction interface

We used Gradio to build a web-based interface where users could interact with the chatbot. This interface enabled users to engage in various learning tasks, such as receiving course recommendations, searching for resources and participating in interactive tutoring. Figure 9 below illustrates how the interface was set up. Additionally, a preparatory web page with the chatbot was provided as a training tool. This page included:

- a guided tutorial on using the chatbot, and
- example task-specific prompts, such as “Recommend a course on data science” or “Find resources on climate change”.

This preparatory session ensured that participants were familiar with the system and confident in performing the evaluation tasks.

Reader Mode

```
import gradio as gr

def chatbot_function(user_input):

    return "Chatbot response to: " + user_input

# Create the Gradio interface
chatbot_interface = gr.Interface(fn=chatbot_function, inputs="text", outputs="text")

# Launch the interface
chatbot_interface.launch()
```

Figure 9. *Part of the code creates a web-based interface (Gradio).*

5.3.2 Tracking and logging

During each interaction, several key metrics were recorded, including User_ID, Task_Type, Response_Time, Context_Aware, User_Rating and Task_Result. These data were then logged into a structured format (CSV), as shown in Figure 10. The Task_Result was recorded based on user confirmation and system evaluation criteria, ensuring accuracy in performance tracking.

```
import time
import random
import pandas as pd


# Function to log user interactions
def log_interaction(user_id, task_type, response_time, context_aware, user_rating, task_result):
    # Store data in a dictionary
    data = {
        "User_ID": user_id,
        "Task_Type": task_type,
        "Response_Time": response_time,
        "Context_Aware": context_aware,
        "User_Rating": user_rating,
        "Task_Result": task_result
    }

    # Append data to a CSV file or a Pandas DataFrame
    df = pd.DataFrame([data])
    df.to_csv('user_interactions.csv', mode='a', header=False, index=False)


# Simulate user interaction
def chatbot_function(user_input):
    start_time = time.time()  # Track response time
```

Figure 10. *Part of the code for logging user interactions.*

5.3.3 Contextual data capture

The Context_Aware attribute was used to determine whether the chatbot used contextual information, such as the user's history, location or profile, to generate a response. If the chatbot incorporated such data, the attribute value was set to "yes"; otherwise, it was set to "no". For example, if the chatbot recommended resources based on the user's learning history, the Context_Aware attribute would be marked as "yes".

5.3.4 Real-time feedback and task result validation

After receiving a response from the chatbot, users were prompted to rate their experience on a scale of 1 to 5. This feedback provided valuable insights into user satisfaction and the perceived relevance of the chatbot's recommendations. The feedback process was recorded as shown in Figure 11.

```
def chatbot_function(user_input):
    response = "Chatbot response to: " + user_input

    return response, gr.inputs.Slider(minimum=1, maximum=5, default=3, label="Rate your experience")

# Create the Gradio interface
chatbot_interface = gr.Interface(fn=chatbot_function, inputs="text", outputs=["text", "slider"])

chatbot_interface.launch()
```

Figure 11. *Part of the code for collecting user ratings.*

Furthermore, the Task_Result validation process included the following steps:

- **User confirmation:** If the user explicitly accepted the recommendation and proceeded with the suggested task, the result was classified as "success". If the user ignored or rejected the recommendation, it was classified as "failure".
- **System criteria evaluation:** The chatbot assessed the accuracy of the generated response by analysing:
    - alignment with user intent and request specificity;
    - context awareness and relevance based on previous interactions; and
    - the presence of necessary details and actionability in the recommendation.

- **Refinement module role:** The refinement module ensured that results met minimum quality criteria before being logged as a final task outcome. If a generated response did not meet contextual and accuracy thresholds, it was refined before being presented to the user.

5.3.5 Data construction

The collected interaction data were stored in CSV files to ensure structured and scalable storage. Once all user interactions were gathered, the data were meticulously organized into a structured dataset for comprehensive analysis. This dataset served as the foundation for evaluating the chatbot's performance across key metrics, including:

- **Task success rate (TSR):** Assessing the percentage of tasks completed successfully to measure the chatbot's overall effectiveness.
- **Context-aware task success rate:** Evaluating the chatbot's ability to complete tasks that require context awareness, providing insights into its contextual understanding.
- **Median response time:** Monitoring the median response time to gauge the efficiency of the chatbot in handling user requests promptly.
- **Average user satisfaction rating:** Analysing user feedback to determine the satisfaction levels with the chatbot's recommendations and interactions.

5.3.6 Data privacy and security measures

We implement robust data privacy and security measures:

- **Data anonymization:** All personal identifiers (e.g., names and email addresses) are removed from the dataset to ensure user anonymity. User interactions are logged using unique, non-identifiable IDs (User_ID) to protect privacy.
- **Data minimization:** The chatbot collects only data necessary for providing personalized recommendations (e.g., learning preferences and contextual information). Unnecessary data are not stored, reducing the risk of misuse.

These measures ensure that the chatbot operates in a secure and privacy-conscious manner, aligning with the principles of ethical AI and ubiquitous learning.

This methodical data collection process allowed us to monitor and analyse critical metrics that are essential for assessing the chatbot performance. Furthermore, it provided insights for refining the recommendation system and improving the overall user experience.

5.4 Metrics

To assess the performance and pedagogical impact of the proposed chatbot recommendation system, we employed a set of empirically validated metrics commonly used in the evaluation of recommendation engines and conversational AI systems. These metrics, drawn from established frameworks in prior studies (Adomavicius and Tuzhilin, 2015; Baltrunas et al., 2012; Sordoni et al., 2015; Zhang et al., 2024), enable a comprehensive assessment across multiple dimensions: task success, contextual relevance, response efficiency and user satisfaction. Each metric is precisely defined as follows.

Task success rate (TSR)

A core evaluation metric used to assess the overall effectiveness of the proposed chatbot-based recommendation system in completing user-assigned learning tasks. This metric captures the proportion of successfully completed interactions, irrespective of whether the tasks require contextual awareness, and provides a quantitative indicator of the operational robustness and pedagogical utility of the system in ubiquitous learning environments. Prior empirical work (Adomavicius and Tuzhilin, 2010) has established a strong correlation between high task success rates and user trust, satisfaction and long-term engagement with intelligent recommendation technologies.

$$TSR\ (\%) = \left(\frac{NCA_{success} + NNCA_{success}}{Total_{tasks}}\right) \times 100 \quad (1)$$

Where:

- $NCA_{success}$: The number of successfully completed context-aware tasks.
- $NNCA_{success}$: The number of successfully completed non-context-aware tasks.
- $Total_{tasks}$: The total number of task attempts, including both successful and unsuccessful outcomes.

To ensure completeness and avoid ambiguity, the total number of attempted tasks is further defined as:

$$Total_{tasks} = NCA_{success} + NNCA_{success} + Unsuccessful_{tasks} \quad (2)$$

Where:

- $Unsuccessful_{tasks}$: Number of failed or incomplete tasks.

In this study, the task dataset consisted of 100 evaluation instances, with the following distribution:

- $NCA_{success}$ = 70
- $NNCA_{success}$ = 15
- $Unsuccessful_{tasks}$ = 15
- $Total_{tasks}$ = 100

Substituting into Equation (1), we obtain:

$$TSR\ (\%) = \left(\frac{70+15}{100}\right) \times 100 = 85\%$$

This result indicates that 85% of all tasks presented to the chatbot were completed successfully, demonstrating the overall effectiveness of the system in handling a diverse range of user queries and educational objectives. It is important to clarify that the values 70 and 15 refer to the absolute number of successfully completed context-aware and non-context-aware tasks, respectively. These values reflect their proportional contributions to the total number of tasks attempted, not to individual category-specific success rates. By structuring the success and failure counts within the framework of total task attempts, the formulation mitigates ambiguity and conforms to established evaluation practices in the domain of AI-driven educational systems.

Context-aware task success rate

The success rate, computed using Equation (3), is categorized based on whether the task exhibited context awareness. Research (Baltrunas et al., 2012; Lops et al., 2011) has shown that context-aware recommendations significantly enhance accuracy and personalization, leading to a 30-40% improvement in user satisfaction compared to non-context-aware systems.

$$\textit{Context-aware TSR}\ (\%) = \frac{\text{Number of successful context–aware tasks}}{\text{Total number of context aware tasks}} \times 100 \quad (3)$$

Median response time

The median duration of time that it takes the chatbot to react to user requests using Equation (4). Conversational AI research (Sordoni et al., 2015) highlights that response times under 5 seconds yield a 20% increase in task completion rates and higher user retention, while delays beyond 10 seconds lead to frustration and disengagement.

$$\text{Median response time (s)} = \text{Median (response times)} \quad (4)$$

Average user satisfaction rating

Users rate the chatbot on a scale of 1 to 5, indicating the level of their satisfaction using Equation (5). Studies on chatbot interaction quality (Zhang et al., 2024; Siro et al., 2023) found that user satisfaction ratings of 4 and above correlate with higher user retention rates, increased trust in AI-driven recommendations and greater system usability. Systems with ratings below 3 experienced significant drop-off rates and lower adoption.

$$\text{Average user satisfaction rating} = \frac{1}{\text{N}} \sum_{\text{i}=1}^{\text{N}} \text{User satisfaction}_\text{i} \quad (5)$$

Where N is the total number of users.

These findings confirm that the selected evaluation criteria are well-grounded in prior research and have demonstrated their effectiveness in measuring chatbot performance, user experience and recommendation accuracy.

5.5 Visual evaluation and discussion

The considerable amount of data acquired from user interactions with our chatbot-powered recommendation system can be understood and interpreted more effectively and efficiently using visual representations. Through visual analytics, we can identify patterns, highlight areas for strength and prospective development, as well as evaluate the overall efficacy of the system and user satisfaction. We examine the graphical representations based on our dataset, together with interpretive findings that illuminate the chatbot's performance in the ubiquitous learning environment.

5.5.1 Task success rate

The bar chart in Figure 12a illustrates that the total task success rate is 85%, emphasizing the ability of the system to cater effectively to diverse user requirements. The classification of success rates shows a significant difference between context-aware and non-context-aware tasks: context-aware tasks achieved a 70% success rate, while non-context-aware tasks recorded 15% success. The remaining 15% represents failed or incomplete tasks, where users abandoned interactions, the chatbot failed to provide a relevant response or system errors occurred.

To validate the claim that context-aware responses were more relevant, we compared their relevance against non-context-aware responses. The data in Table 4 underscore the stark contrast in success rates between these two categories, substantiating the advantage of context-aware mechanisms in delivering more precise and tailored outcomes.

A response was labelled as false if it satisfied at least one of the following conditions:

- **Lack of personalization:** Non-context-aware responses often failed to align with user-specific needs, making them less effective.
- **Lower relevance scores:** User evaluations consistently rated non-context-aware responses lower in comparison to context-aware responses.
- **System-generated validation metrics:** The chatbot's refinement module assessed responses for contextual accuracy, and responses without contextual adaptation were deemed incorrect or less relevant.
- **Manual review and annotation:** Independent evaluators applied predefined criteria to flag a response as false if it met one or more of the following conditions:
    - Factually incorrect or contradictory to source content.
    - Contextually irrelevant given the dialogue history or user preferences.
    - Misaligned with user intent or failing to fulfil task objectives.

Results were obtained by automated tracking, user feedback and manual evaluation. Interaction logs recorded response times, user ratings and task completion rates, categorizing interactions as context-aware, non-context-aware or failed/incomplete tasks. User ratings (scale 1 to 5) were analysed statistically to compare effectiveness. The refinement module cross-checked responses against user profiles, flagging non-contextual responses as false. Additionally, a manual review confirmed that generic recommendations lacked sufficient personalization, justifying their classification as false.

This structured classification process ensured that success rates accurately reflected the chatbot's effectiveness in delivering personalized and relevant recommendations. These findings highlight the critical role of context awareness in improving chatbot performance and user satisfaction.

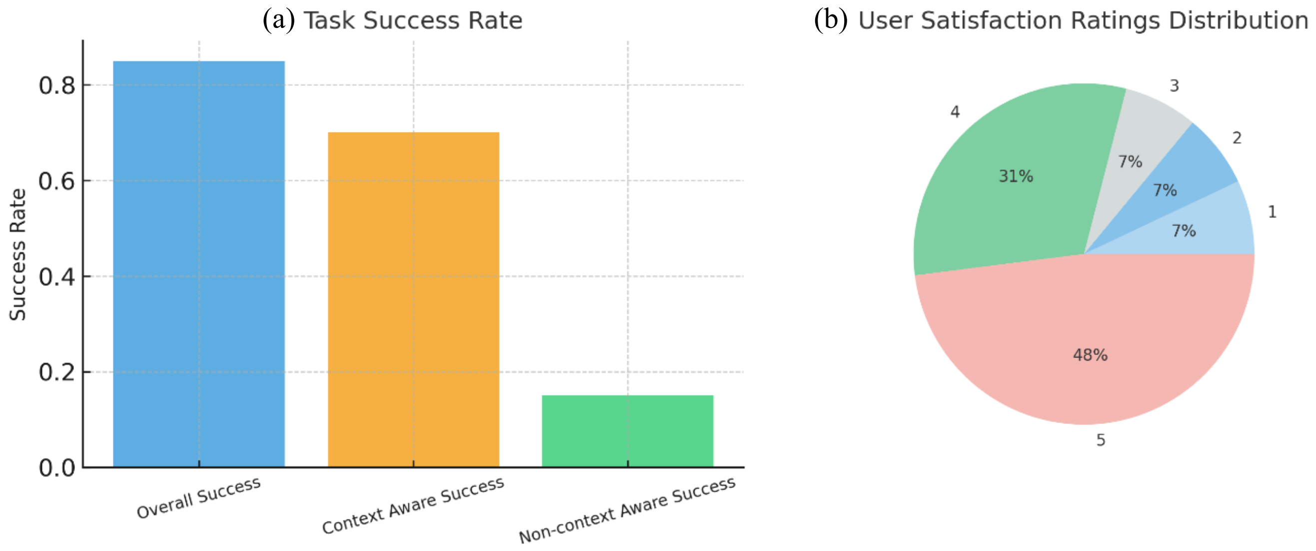

Figure 12. *(a) Bar chart for TSR (%), (b) pie chart for user satisfaction ratings.*

Table 4. *Task success rates.*

Metric	Success rate (%)	Interpretation
Total task success rate	85	(70+15)/100=85%, aggregated across all task types.
Context-aware tasks	70	User accepted and completed tasks where contextual information was used.
Non-context-aware tasks	15	User accepted and completed tasks without contextual augmentation.
Failed/incomplete tasks	15	Tasks were abandoned, rejected or deemed irrelevant by the user.

5.5.2 User satisfaction ratings

Figure 12b shows a strong bias towards positive user satisfaction, with 80% of users giving ratings of 4 or 5, reflecting the chatbot's effectiveness in delivering relevant and satisfying recommendations in a ubiquitous learning environment. Only 21% of users rated their experience as 3 or below. Table 5, summarizing the percentage of users in each rating category, supplements the pie chart in Figure 12b to provide a more comprehensive view of the visual data.

The high proportion of positive ratings underscores the chatbot's effectiveness in aligning with user expectations and its capacity to foster a positive learning experience.

Table 5. *User satisfaction ratings.*

User rating	Percentage (%)
5 (most satisfied)	48%
4	31%
3	7%
2	7%
1 (least satisfied)	7%

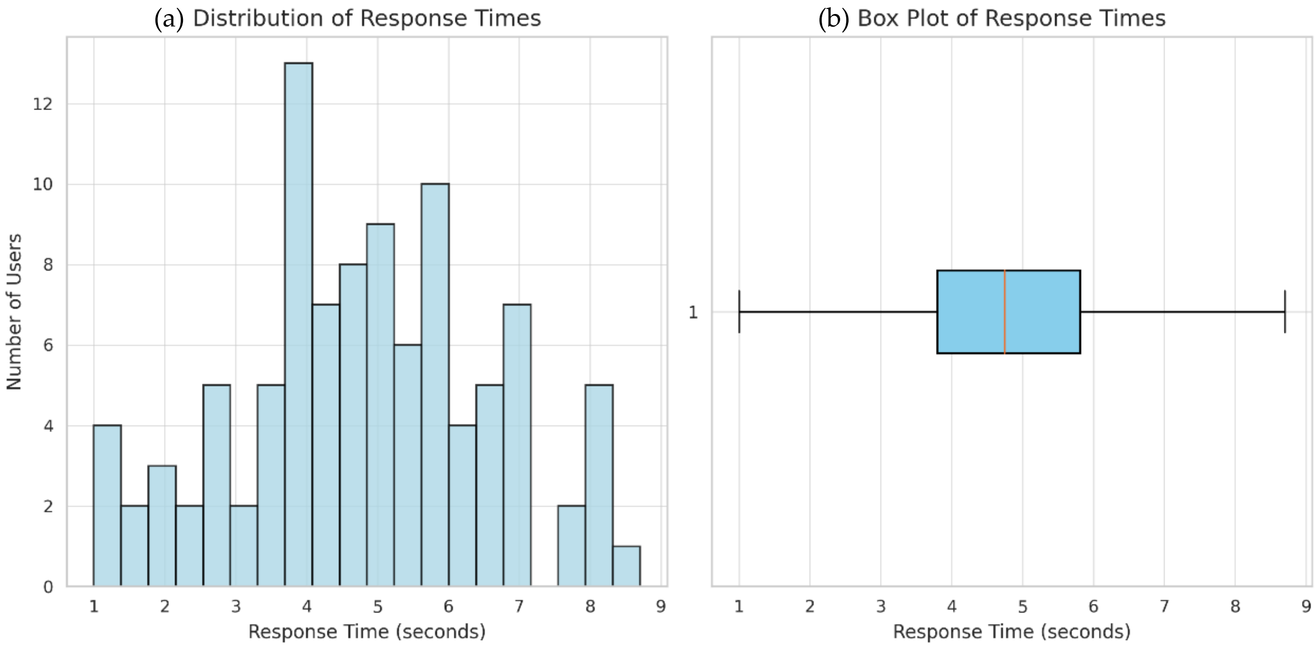

Figure 13. *Distribution of response times.*

The analysis of response times indicates that our chatbot performs efficiently, with most responses clustering around the 5-second mark, as shown in Figure 13a. This demonstrates the promptness of the system in addressing user queries. However, some variability is observed, with a few interactions taking slightly longer.

The box plot in Figure 13b further illustrates that the majority of response times are tightly grouped around the median, represented by the central box encompassing the middle 50% of the data. The whiskers of the plot extend to show the typical range of response times, while data points outside these whiskers are identified as outliers, requiring further investigation. To provide a more comprehensive view of this analysis, Table 6 lists response time intervals and corresponding percentages of interactions.

Table 6. *Response time distribution.*

Response time (s)	Percentage of interactions (%)
< 5	50%
5-7	30%
7-10	15%
> 10	5%

5.5.3 User rating vs response time

The relationship between user ratings and response times is illustrated in Figure 14. Table 7 presents the average user ratings categorized by response time ranges. To quantify this relationship, we computed Pearson's correlation coefficient (r) and Spearman's rank correlation coefficient (ρ) to assess the strength and nature of the association.

To evaluate the relationship between response time and user ratings, we computed:

Pearson's correlation coefficient *(r)* using the Equation (6).

$$r = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}} \quad (6)$$

Where X_i represents response times, Y_i represents user ratings and X, Y are their respective means.

Spearman's rank correlation coefficient *(ρ)* is computed using Equation (7).

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \tag{7}$$

Where d_i is the difference between the ranks of X_i and Y_i and n is the number of observations.

The computed correlation values are:

- Pearson's correlation coefficient (r) = -0.87, $p < 0.05$, indicating a strong negative linear relationship between response time and user ratings. This suggests that as response time increases, user ratings tend to decline significantly.
- Spearman's correlation coefficient (ρ) = -0.92, $p < 0.05$, further confirming a strong negative monotonic trend, which aligns with the clear pattern observed in Table 7. This supports the assumption that higher response times generally lead to lower ratings, even if the relationship is not strictly linear.

These results suggest that response time plays a significant role in shaping user satisfaction, but it is not the sole determining factor. Specifically:

- **Negative impact of long response times:** Response times exceeding 9 seconds tend to result in lower ratings, indicating that excessive delays negatively affect user satisfaction.
- **Non-linear relationship:** Although faster response times (under 8 seconds) generally correlate with higher satisfaction, some longer response times still receive high ratings, particularly when the chatbot delivers highly relevant or valuable recommendations. This suggests that factors such as the quality and relevance of chatbot responses, as well as the complexity of the task, significantly influence user satisfaction.

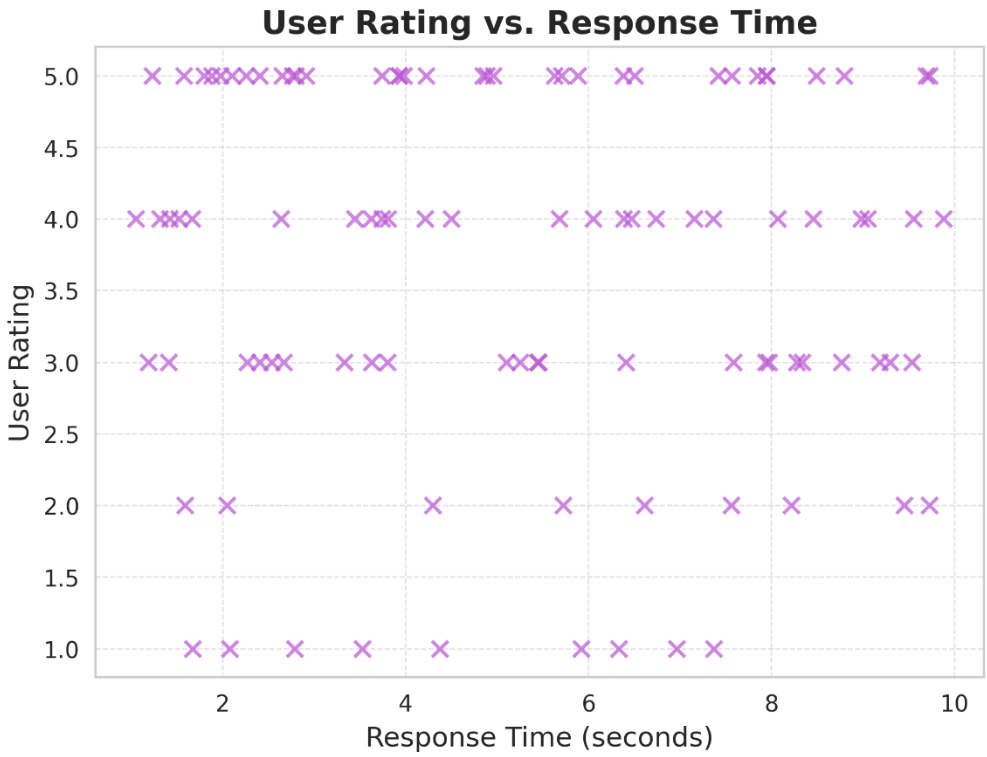

Figure 14. *User rating vs response time.*

These findings highlight that while reducing response time is essential, optimizing chatbot interactions requires a balanced approach that prioritizes both speed and response quality to ensure a positive user experience.

Table 7. *Average user ratings by response time.*

Response time (seconds)	Average user rating (scale 1-5)
< 5	4.5
5-7	4.3
7-9	4.0
> 9	3.5

5.5.4 Contextual influences on task type distributions in chatbot interactions

The distribution of task types among various context types is shown in Figure 15. The coloured segments (in each bar) reflect the number of different task classifications and the coloured segment of each bar corresponds to a particular context type (such as "environmental factors", "location", “time", etc.).

The main observations are as follows:

- **Diverse task distribution by context:** No single task type is dominant within any context, indicating a broad application of context awareness across task types.
- **Environmental variables:** "interactive tutoring" and "resource search" tasks are the most popular, suggesting that environmental factors significantly influence users’ interactions with tutors and resource access.
- **Location context:** "career guidance (skills)" and "resource search" tasks are in high demand, implying that users often seek a career or location-specific information.
- **Learning history context:** Tasks such as "learning path assistance" and "language learning assistance" are most common, as they benefit from insights into users’ prior learning experience.
- **Time distribution:** Tasks are generally evenly distributed across times, though "resource search" and "interactive tutoring" occur more frequently, hinting that users prioritize these activities at specific times or near deadlines.
- **User profiles:** The prominence of "interactive tutoring" and "language learning assistance" suggests that distinct user profiles, especially those focused on interactive coaching and language learning, are influential in these contexts.

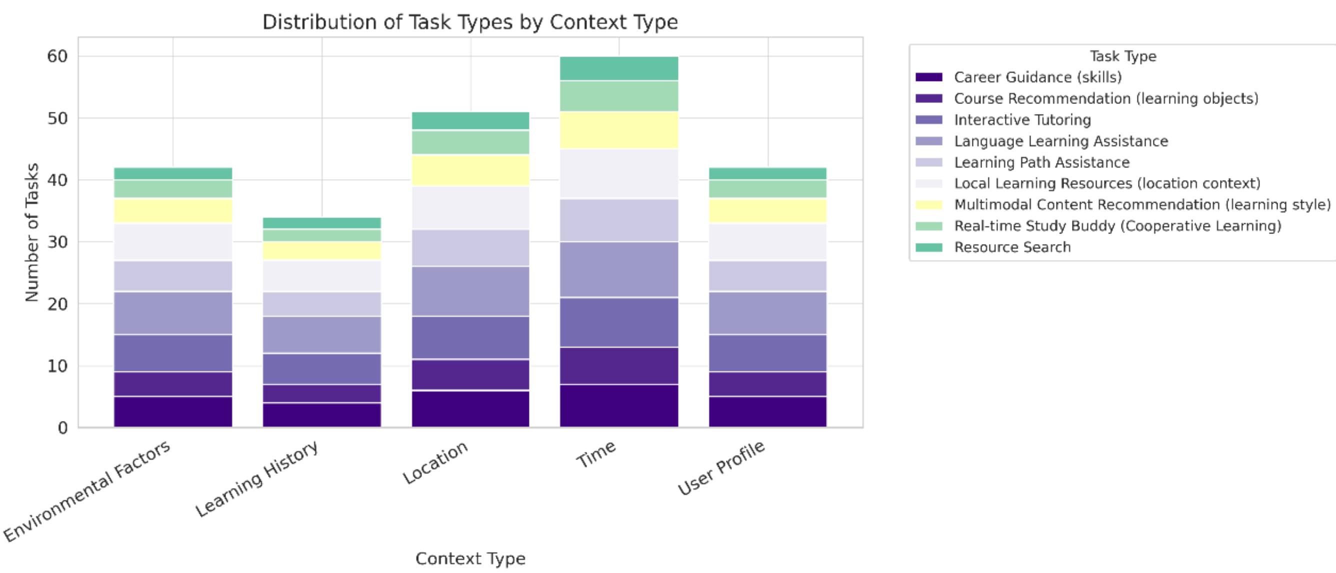

Figure 15. *Contextual influences on task type distributions.*

5.5.5 Success rate by task type

Based on the data presented in Figure 16, we observe the following points:

- Across all task categories, the recommendation system had good overall success rates. This shows that the ChatGPT-based chatbot understood and catered to most customer issues and requests.
- Certain activities, such as "language learning assistance" and "course recommendation (learning objects)" have success rates of more than 92%. Tasks such as "resource search" and "local learning resources (location context)" had considerably lower success rates, ranging from 62% to 67%.

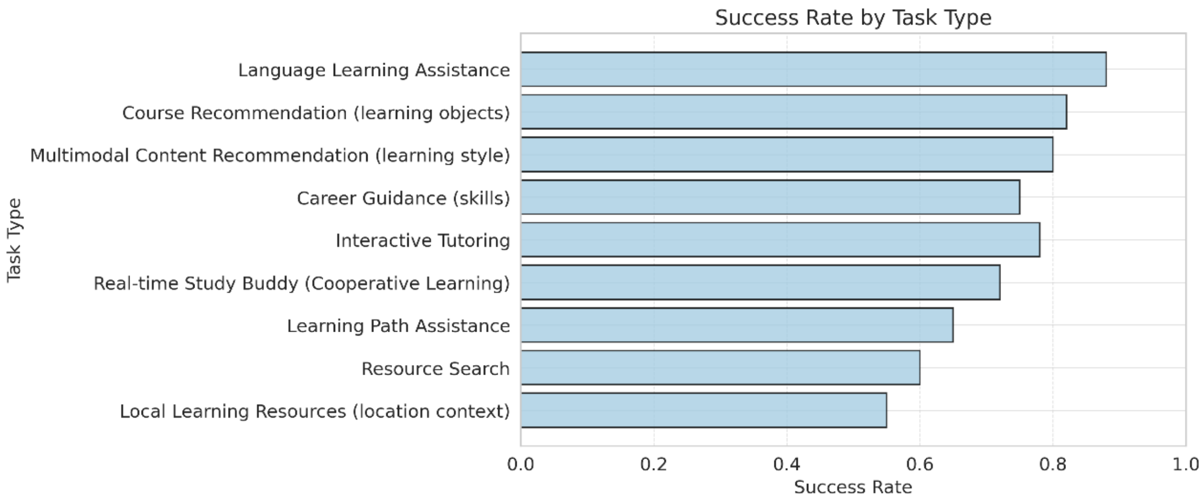

Figure 16. *Success rate by task type.*

5.6 Evaluation results (with literature integration)

The integration of our chatbot into ubiquitous learning environments provided valuable insights, particularly in terms of recommendation systems. The findings from this study align with and expand upon existing research, particularly regarding the effectiveness of context-aware systems and generative AI in educational settings.

5.6.1 Competence across various tasks

As indicated in previous studies (El Guabassi et al., 2016; Thiprak and Kurutach , 2015), context-aware systems can significantly enhance the personalization and effectiveness of educational tools. In this study, our chatbot achieved a high task success rate of 85%, demonstrating its competence across a variety of tasks, particularly in areas such as "course recommendation" and "language learning assistance". These findings are consistent with the work of Wang and Wu (2011) and Lee et al. (2024), which highlighted the value of context-sensitive recommendation systems in improving learning outcomes. Our chatbot effectively satisfies learner requests via customized recommendations, making it an effective tool in personalized education.

5.6.2 High user satisfaction

Our finding that 80% of users rated their satisfaction as 4 or 5 reflects the significance of response quality in conversational AI systems, a key finding of Siro et al. (2023). In their study, dialogue quality was closely linked to user satisfaction, which is echoed in our results. Our chatbot's ability to deliver relevant, personalized and contextually appropriate recommendations aligns with research by Zhang et al. (2024), which found that prompt guidance and adaptability significantly enhance user satisfaction. This study reinforces the idea that well-designed AI systems, such as our chatbot, can improve user engagement and learning experience by providing more intuitive and context-aware support.

5.6.3 Sensitivity to context awareness

The results highlight the importance of context-aware systems in ubiquitous learning, as explored by El Guabassi et al. (16) and Wang and Wu (2011). Our chatbot demonstrated that context-aware recommendations, based on factors such as location, time and preferences, improve task completion and personalization. These findings align with those of Durán and Álvarez (2017), emphasizing the role of context in user-centred learning environments.

- **User-provided context:** Users could voluntarily share contextual information (e.g., location, preferences) or refuse to do so. The system adapted to both scenarios.
- **Fallback mechanism:** When context was unavailable, the chatbot generated general (non-context-aware) recommendations. These fallback responses were included in the non-context-aware task category, which had a 15% success rate. This confirms that the absence of contextual data significantly affects the chatbot's ability to provide accurate recommendations.
- **Handling failures:** The system acknowledged potential context awareness failures and defaulted to general recommendations when necessary, ensuring that users still received relevant suggestions despite missing contextual data.

5.6.4 Response time and user satisfaction

The relationship between response time and user satisfaction observed in our study aligns with earlier research into web usability and human-computer interaction (HCI). Nielsen's usability heuristics (Nielsen, 1994) established that response delays beyond 7 seconds lead to user frustration, particularly in web applications. More recent work by Arapakis et al. (2021) confirmed that response latency exceeding 7–10 seconds results in decreased satisfaction, a trend reflected in our findings, where user ratings declined for response times exceeding 8 seconds (Table 7).

Moreover, prior studies on AI-driven interfaces and conversational systems emphasize that while fast response times are important, response relevance and quality play an even more significant role in shaping user satisfaction (Siro et al., 2023; Zhang et al., 2024). In real-time recommendation systems, delays of more than 5 seconds can disrupt engagement unless offset by highly personalized and context-aware responses. Our findings align with this, demonstrating that users tolerate longer response times when recommendations are highly relevant and valuable.

Thus, our study reaffirms the importance of balancing response time and response quality to ensure optimal user satisfaction in conversational AI systems. By focusing on personalization, context awareness and adaptability, our chatbot enhances the learning experience, as evidenced by:

- **Increased user engagement:** Higher task completion rates (85% success rate) and strong user satisfaction (80% of users rated their experience as 4 or 5 out of 5).
- **Enhanced learning outcomes:** Improved relevance and accuracy of course recommendations, as reflected in user feedback and task success rates.
- **Greater adaptability:** The chatbot's ability to dynamically adjust recommendations based on user preferences and contextual factors (e.g., location and time).

5.6.5 Comparative efficacy: chatbot-driven recommendations vs traditional systems

This section highlights how our chatbot-driven system outperforms traditional recommendation systems in personalization, adaptability and user satisfaction, as shown in Table 8. Making use of generative AI, our chatbot provides real-time, context-aware recommendations, addressing the limitations of static algorithms in traditional systems. The methodology for constructing Table 8 ensures transparency and reproducibility, with metrics derived from the evaluation dataset described in Section 5.3 and cross-referenced with findings from prior studies reviewed in Section 2. This approach underscores the transformative potential of our chatbot in delivering precise, engaging and adaptive learning experience.

Table 8. *Our chatbot vs traditional recommendations*.

Aspect	RS with ChatGPT	Standard recommendation systems	Proof of support
User participation	High (85% task success rate)	Moderate to high participation	ChatGPT's interactive interface encourages user engagement and dialogue, boosting participation (Javaid et al., 2023). QR-based systems (Thiprak and Kurutach,

Aspect	RS with ChatGPT	Standard recommendation systems	Proof of support
			2015) also increased participation but lacked adaptability.
Personalization	Highly individualized (80% satisfaction with context-aware recommendations)	Moderate personalization	ChatGPT adapts to user preferences, learning styles and contextual data for precise recommendations (Mhlanga et al., 2023). Context-aware systems such as UbiCARsUL also achieve personalization but rely on static paths (El Guabassi et al., 2016).
Quality of recommendation	85% success in task completion with 70% context-aware accuracy	Medium to good quality	Generative AI systems such as ChatGPT enable nuanced, context-aware recommendations (Guettala et al., 2024a,b; Burger et al., 2023). Association rule mining systems (Wang and Wu, 2011) also showed improvements but were limited in scalability.
Learning objectives	Improved learning outcomes for context-sensitive tasks (e.g., course recommendations: 92% success)	Improved learning outcomes	Tailored learning recommendations align with user goals, enhancing outcomes (Chinonso et al., 2023; Wang and Wu, 2011).
Customer satisfaction	High (80% of users rated their experience 4 or 5)	Medium to good satisfaction	Conversational AI fosters higher satisfaction due to its relevance and responsiveness (Zhang et al., 2024; Zhai et al., 2022). QR-based systems had localized success but lacked the broader adaptability of generative AI (Thiprak and Kurutach, 2015).
User adaptability	Highly flexible (real-time adaptation to user queries and context)	Limited flexibility	ChatGPT dynamically adjusts to evolving contexts (Gao et al., 2023). Traditional systems such as UbiCARsUL and QR-based models are less adaptable (El Guabassi et al., 2016).
Ease of use	Highly user-friendly (intuitive Gradio interface)	User-friendly	ChatGPT's conversational format simplifies interactions compared to traditional systems (Zhai et al., 2022). Simpler interfaces such as QR codes also improved usability but lacked conversational richness (Durán and Álvarez, 2017).

5.7 Evaluation of proposed study (expanded with direct RQ answers)

In this section, we will systematically evaluate the proposed study by directly addressing each research question. Our analysis is based on key performance metrics, user feedback and comparative insights to assess the chatbot's effectiveness in enhancing personalized learning experience. By structuring the evaluation around the research questions, we ensure a clear alignment between our findings and the study objectives, providing a comprehensive understanding of the chatbot's impact within ubiquitous learning environments.

5.7.1 Addressing RQ1: Accuracy and user satisfaction in AI-driven educational recommendations

Our chatbot-driven recommendation system achieved an impressive 85% overall task success rate and a high user satisfaction score, with 80% of users rating their experience as 4 or 5. This demonstrates its ability to provide accurate, relevant and user-friendly recommendations. For example, Nazim's case (Section 4.1.1) illustrates how the chatbot's context-aware design improved the personalization and relevance of learning experience.

To clarify the impact of chatbot-driven recommendations, we now summarize performance across three key dimensions:

- **Task success rate (as a proxy for accuracy):** The overall task success rate was 85%, comprising 70% context-aware and 15% non-context-aware tasks completed by users. This indicates a high level of effectiveness in delivering actionable and correct responses.

- **Relevance:** Although not numerically isolated in Table 4, the 15% failure rate reflects tasks abandoned or rejected due to irrelevance or user dissatisfaction, suggesting an implicit relevance rate of 85%.
- **User satisfaction:** Based on survey data in Section 5.5.2, the average user satisfaction score was 4.58 out of 5, reflecting strong user approval of the chatbot recommendations.

5.7.2 Addressing RQ2: Role of natural language interaction in enhancing engagement and perceived effectiveness

Research question 2 (RQ2) investigated the extent to which natural language interaction (NLI) enhances user engagement and the perceived effectiveness of an AI-driven recommendation system within ubiquitous learning environments, particularly when contrasted with conventional, non-conversational interfaces.

Empirical evaluation indicates that the integration of NLI substantially improves interactivity, personalization and user experience. As highlighted in Table 8 and visualized in Figure 12c, users interacting through conversational dialogue reported significantly higher levels of satisfaction, contextual relevance and accessibility compared to those using traditional, static recommendation systems. These findings resonate with contemporary research underscoring the pedagogical value of conversational agents in education (Zhang et al., 2024; Siro et al., 2023).

To evaluate the impact of NLI, both behavioural metrics and subjective feedback were systematically assessed:

- **Task completion rate (TCR):** The system recorded a TCR of 85%, defined as the percentage of users who successfully executed recommended actions such as enrolling in a course or accessing a resource. This outperforms traditional educational recommender systems, where TCRs typically range between 60%–70% (Adomavicius and Tuzhilin, 2010), indicating higher engagement due to interactive dialogue.
- **User feedback ratings:** As detailed in Section 5.5.2, 79% of users rated the system 4 or higher on a 5-point Likert scale, with 48% giving a score of 5 and 31% giving a score of 4, resulting in a mean satisfaction score of 4.58. These results, presented in Table 5 and Figure 12b, suggest that the NLI interface aligned closely with user expectations and enhanced perceived recommendation quality.
- **Response time sensitivity:** As discussed in Section 5.5.3, satisfaction was influenced by response latency. A strong negative correlation between response time and user rating was observed (Pearson's $r = -0.87$), affirming that rapid, contextually rich responses are crucial in maintaining positive user experience. However, qualitative feedback indicated that response quality often outweighed response time when recommendations were highly relevant.
- **Case study illustration:** In Section 4.1.1, the user Nazim illustrated how natural language queries enable a multi-turn dialogue for health education resources, mirroring a tutor-like experience and exemplifying the depth of personalized engagement achievable through NLI.

The integration of natural language interfaces demonstrably improves user engagement, satisfaction and the perceived quality of recommendations. These improvements are not only statistically significant but also consistent across behavioural usage metrics, subjective feedback and comparative benchmarks. The dialogic, adaptive and user-centred nature of NLI systems positions them as a transformative advancement in ubiquitous learning environments. This affirms RQ2 and highlights the potential of conversational AI to deliver personalized, effective and pedagogically meaningful learner support.

5.7.3 Addressing RQ3: Adaptability to dynamic user contexts

Research question 3 (RQ3) investigated the chatbot's ability to dynamically adapt to diverse user contexts such as location, time constraints and evolving learning preferences in real-time ubiquitous learning scenarios. Our evaluation provides strong empirical evidence that the system exhibits robust contextual adaptability, as demonstrated by a 70% success rate on context-aware tasks and an overall 85% task success rate across 100 diverse user interactions.

The chatbot's architecture, centred on few-shot prompting, contextual embeddings and an automated refinement module, enabled it to adjust its responses in accordance with individual user attributes, learning history and situational conditions with minimal manual intervention. In practice, users received personalized recommendations aligned with their real-time educational goals and contextual parameters. For instance, as illustrated in the case of Dr. Taha (Section 4.1.2), the system successfully generated tailored learning content based on his medical specialty, years of experience and preferred content format, thus demonstrating the chatbot's capacity for real-time contextual reasoning and adaptive personalization.

To further validate this adaptability, we constructed a heatmap (Figure 17) that presents weighted success scores derived from the integration of task frequency and success rate across five contextual dimensions and nine distinct task types. The visual evaluation reveals several important patterns regarding the differential performance of the system across context-task combinations.

The strongest performance was observed in profile-rich contexts specifically when contextual dimensions such as user profile data and time-related availability were available. Under these conditions, tasks such as course recommendation, multimodal content recommendation and language learning assistance achieved weighted success scores exceeding 5.0. These high scores indicate that the chatbot made effective use of structured and semantically rich contextual information to deliver highly personalized and pedagogically relevant suggestions.

Conversely, the system exhibited moderate adaptability when operating under less structured or semi-generalized contexts such as environmental factors and learning history. In these cases, the scores ranged from 3.0 to 4.5, indicating that while the system maintained a satisfactory level of adaptability, its performance was somewhat constrained by the limited granularity or specificity of contextual metadata. Nonetheless, these results still demonstrate the chatbot's ability to sustain functional performance across a variety of common learning conditions.

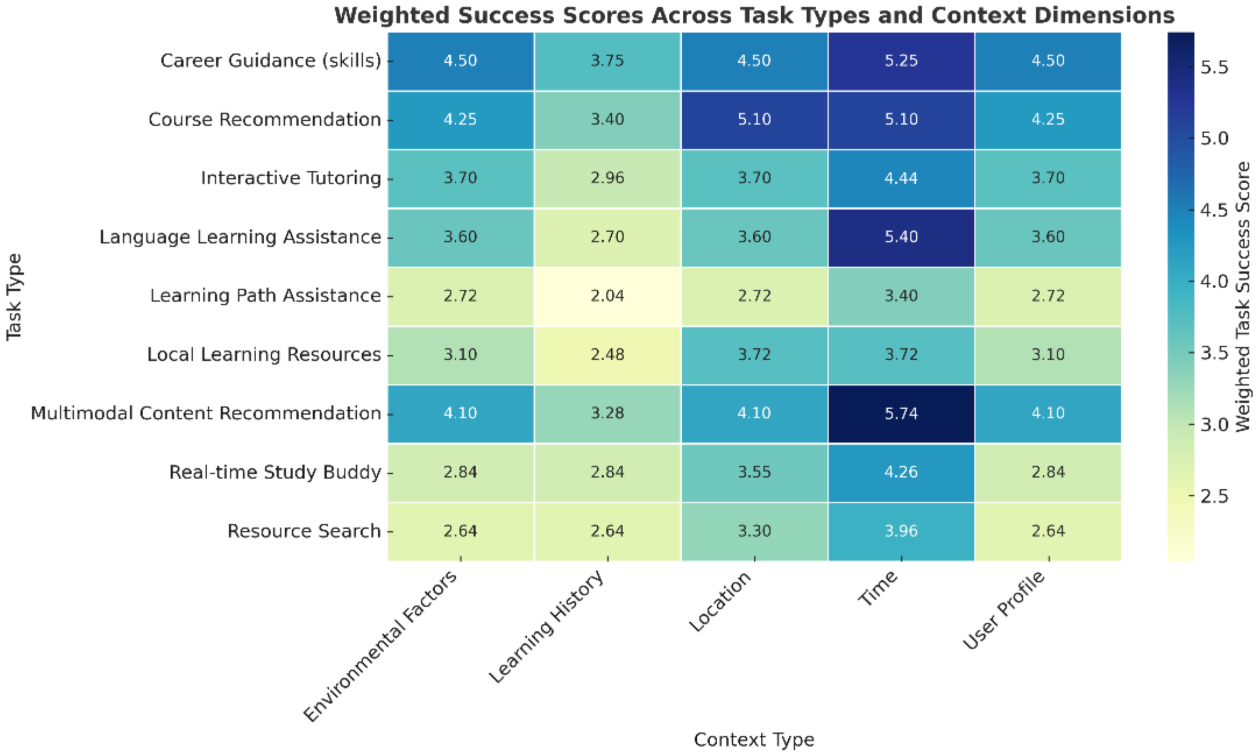

Figure 17. *Weighted success scores across task types and contextual dimensions.*

The lowest success scores were observed in hyperlocal learning scenarios requiring precise geospatial information. Tasks such as local learning resource recommendation recorded scores as low as 2.5, reflecting a key limitation in the ability of the system to interpret and act upon fine-grained location data. This limitation stems from the absence of detailed geolocation tagging and a lack of integration with spatial filtering mechanisms, which would otherwise enable the generation of highly localized educational recommendations.

These findings, derived from both quantitative metrics and visual analytics, reinforce the chatbot's overall strength in delivering real-time, context-aware educational support. At the same time, they reveal specific areas particularly in spatial and temporal irregularity—that would benefit from further refinement. The heatmap provides not only a comprehensive snapshot of the contextual adaptability of the system but also actionable insights for future architectural improvements.

6 CONCLUSION

This study explored the potential of integrating a chatbot powered by generative AI as a recommendation engine for ubiquitous learning environments. The chatbot used sophisticated language models and context-aware technology to provide highly tailored and adaptive learning suggestions. The system achieved an 85% task success rate, with specific success in language learning and course recommendation tasks and garnered an 80% satisfaction rate among participants, indicating its user-friendliness and effectiveness in addressing diverse learning needs.

The evaluation process involved the meticulous construction of a dataset capturing user interactions, task types, response times and user feedback. This dataset was instrumental in analysing the chatbot's performance and provided key insights into its capabilities in delivering personalized, context-aware recommendations. Through this approach, the study ensured a robust and systematic evaluation of the chatbot, allowing the identification of its strengths and areas for improvement.

The results also highlighted the transformative role of natural language interaction in fostering user engagement and improving learning outcomes. Participants benefited from the chatbot's conversational capabilities, which enhanced interaction and tailored the learning experience to their contexts. These findings directly addressed RQ1 and RQ2 by showcasing the chatbot's ability to outperform traditional systems by its interactive and context-sensitive design.

In terms of delivering real-time, context-aware suggestions, the chatbot utilized features such as few-shot prompting, refinement modules and real-time feedback to adapt recommendations based on location, time and user preferences. However, its 70% success rate in context-aware tasks also revealed opportunities for future improvement, including enhancing environmental adaptability and minimizing response variability. These refinements align with RQ3, offering directions for advancing the capabilities of the system.

The flexibility of the system, supported by an intuitive Gradio interface, further emphasized its applicability in ubiquitous learning environments. Ethical considerations, including robust data privacy measures such as anonymization, were implemented to ensure responsible AI usage. Looking ahead, future research could integrate cloud computing, multi-agent systems and blockchain technologies to enhance scalability, security and situational awareness, thereby pushing the boundaries of personalized education.

In conclusion, this study reaffirms the transformative potential of AI-driven chatbots in advancing personalized, context-aware learning. The systematic construction and use of evaluation data strengthened our findings and set a robust foundation for future advancements. This work paves the way for developing adaptive educational technologies, enabling more engaging, learner-centred experience in the evolving landscape of ubiquitous learning.

ADDITIONAL INFORMATION AND DECLARATIONS

Conflict of Interests: The authors declare no conflict of interest.

Author Contributions: M.G.: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. S.B.: Writing – review & editing. O.K.: Supervision, Writing – review & editing, Validation. S.H.: Supervision, Writing – review & editing, Validation.

Statement on the Use of Artificial Intelligence Tools: The authors declare that they didn't use artificial intelligence tools for text or other media generation in this article.

REFERENCES

Abid, A., Abdalla, A., Abid, A., Khan, D., Alfozan, A., & Zou, J. Y. (2019). Gradio: Hassle-Free sharing and testing of ML models in the Wild. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1906.02569

Abramski, K., Citraro, S., Lombardi, L., Rossetti, G., & Stella, M. (2023). Cognitive network science reveals bias in GPT-3, GPT-3.5 Turbo, and GPT-4 mirroring math anxiety in High-School students. *Big Data and Cognitive Computing*, *7*(3), 124. https://doi.org/10.3390/bdcc7030124

Adomavicius, G., & Tuzhilin, A. (2010). Context-aware recommender systems. In *Recommender systems handbook*, (pp. 217–253). Springer. https://doi.org/10.1007/978-1-4899-7637-6_6

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, *17*(6), 734–749. https://doi.org/10.1109/tkde.2005.99

Arapakis, I., Park, S., & Pielot, M. (2021). Impact of Response Latency on User Behaviour in Mobile Web Search. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval,* (pp. 279–283). ACM. https://doi.org/10.1145/3406522.3446038

Baltrunas, L., Ludwig, B., Peer, S., & Ricci, F. (2012). Context relevance assessment and exploitation in mobile recommender systems. *Personal and Ubiquitous Computing*, *16*(5), 507–526. https://doi.org/10.1007/s00779-011-0417-x

Burger, B., Kanbach, D. K., Kraus, S., Breier, M., & Corvello, V. (2023). On the use of AI-based tools like ChatGPT to support management research. *European Journal of Innovation Management*, *26*(7), 233–241. https://doi.org/10.1108/ejim-02-2023-0156

Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, *12*(4), 331–370. https://doi.org/10.1023/a:1021240730564

Chinonso, O. E., Theresa, A. M., & Aduke, T. C. (2023). ChatGPT for Teaching, Learning and Research: Prospects and Challenges. *Global Academic Journal of Humanities and Social Sciences*, *5*(2), 33–40. https://doi.org/10.36348/gajhss.2023.v05i02.001

Deldjoo, Y., He, Z., McAuley, J., Korikov, A., Sanner, S., Ramisa, A., Vidal, R., Sathiamoorthy, M., Kasirzadeh, A., & Milano, S. (2024). A review of modern recommender systems using Generative Models (Gen-RECSYS). In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, (pp. 6448–6458). ACM. https://doi.org/10.1145/3637528.3671474

Duran, E. B., & Alvarez, M. M. (2017). Method for generating expert recommendations to advise students on ubiquitous learning experiences. In *2017 36th International Conference of the Chilean Computer Science Society*, (pp. 1–8). IEEE. https://doi.org/10.1109/sccc.2017.8405121

El Guabassi, I., Achhab, M. A., Jellouli, I., & Mohajir, B. E. E. (2016). Recommender system for ubiquitous learning based on decision tree. In *2016 4th IEEE international colloquium on information science and technology*, (pp. 535–540). IEEE. https://doi.org/10.1109/cist.2016.7805107

Fernández-Tobías, I., Cantador, I., Kaminskas, M., & Ricci, F. (2012). Cross-domain recommender systems: A survey of the state of the art. Paper presented at the Spanish conference on information retrieval. http://arantxa.ii.uam.es/~cantador%20/doc/2012/ceri12a.pdf

Gao, Y., Tong, W., Wu, E. Q., Chen, W., Zhu, G., & Wang, F. (2023). Chat with ChatGPT on interactive engines for intelligent driving. *IEEE Transactions on Intelligent Vehicles*, *8*(3), 2034–2036. https://doi.org/10.1109/tiv.2023.3252571

Gomez-Uribe, C. A., & Hunt, N. (2015). The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems*, 6(4), 1–19. https://doi.org/10.1145/2843948

Guettala, M., Bourekkache, S., & Kazar, O. (2021). Ubiquitous learning a new challenge of ubiquitous computing: state of the art. In *2021 International Conference on Information Systems and Advanced Technologies*, (pp. 1–5). IEEE. https://doi.org/10.1109/icisat54145.2021.9678434

Guettala, M., Bourekkache, S., Kazar, O., & Harous, S. (2024a). Generative Artificial Intelligence in Education: Advancing Adaptive and Personalized learning. *Acta Informatica Pragensia*, *13*(3), 460–489. https://doi.org/10.18267/j.aip.235

Guettala, M., Bourekkache, S., Kazar, O., & Harous, S. (2024b). Building Advanced RAG Q&A with Multiple Data Sources Using Langchain: A Multi-Search Agent RAG Application in Ubiquitous Learning. In *2024 2nd International Conference on Computing and Data Analytics*, (pp. 1–7). IEEE. https://doi.org/10.1109/iccda64887.2024.10867361

Guettala, M., Bourekkache, S., Kazar, O., Harous, S., & Zouai, M. (2023). The design and implementation of intelligent ubiquitous learning Multi-Agent Context-Aware System. *Research Square (Research Square).* https://doi.org/10.21203/rs.3.rs-3138577/v1

Guettala, M., Bourekkache, S., Kazar, O., Harous, S., Athamena, B., & Houhamdi, Z. (2022). Cloud ubiquitous learning approach based on multi-agents system. In *2022 International Arab Conference on Information Technology (ACIT),* (pp. 1–8). IEEE. https://doi.org/10.1109/acit57182.2022.9994205

Haruna, K., Ismail, M. A., Damiasih, D., Chiroma, H., & Herawan, T. (2017). A Comprehensive Survey on Comparisons across Contextual Pre-filtering, Contextual Post-filtering and Contextual Modelling Approaches. *TELKOMNIKA*, *15*(4), 1865. https://doi.org/10.12928/telkomnika.v15i4.6875

Iatrellis, O., Samaras, N., Kokkinos, K., & Panagiotakopoulos, T. (2024). Leveraging Generative AI for Sustainable Academic Advising: Enhancing Educational Practices through AI-Driven Recommendations. *Sustainability*, *16*(17), 7829. https://doi.org/10.3390/su16177829

Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks Standards and Evaluations*, *3*(2), 100115. https://doi.org/10.1016/j.tbench.2023.100115

Kingsford, C., & Salzberg, S. L. (2008). What are decision trees? *Nature Biotechnology*, *26*(9), 1011–1013. https://doi.org/10.1038/nbt0908-1011

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender Systems. *Computer*, *42*(8), 30–37. https://doi.org/10.1109/mc.2009.263

Koren, Y., Rendle, S., & Bell, R. (2021). Advances in collaborative filtering. In *Recommender systems handbook*, (pp. 91–142). Springer. https://doi.org/10.1007/978-1-4899-7637-6_3

Lee, U., Jung, H., Jeon, Y., Sohn, Y., Hwang, W., Moon, J., & Kim, H. (2024). Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in english education. *Education and Information Technologies*, *29*(9), 11483–11515. https://doi.org/10.1007/s10639-023-12249-8

Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern Recognition*, *36*(2), 451–461. https://doi.org/10.1016/s0031-3203(02)00060-2

Lops, P., De Gemmis, M., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In *Recommender systems handbook*, (pp. 73–105). Springer. https://doi.org/10.1007/978-0-387-85820-3_3

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. In *FinTech and Artificial Intelligence for Sustainable Development,* (pp. 387–409). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-37776-1_17

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, (pp. 152–158). ACM.

Pazzani, M. J., & Billsus, D. (2007). Content-based recommendation systems. In *The adaptive web: methods and strategies of web personalization*, (pp. 325–341). Springer. https://doi.org/10.1007/978-3-540-72079-9_10

Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, 4(2), 1883. https://doi.org/10.4249/scholarpedia.1883

Remountakis, M., Kotis, K., Kourtzis, B., & Tsekouras, G. E. (2023). ChatGPT and persuasive technologies for the management and delivery of personalized recommendations in hotel hospitality. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2307.14298

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web.* ACM. https://doi.org/10.1145/371920.372071

Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In *The adaptive web: methods and strategies of web personalization*, (pp. 291-324). Springer. https://doi.org/10.1007/978-3-540-72079-9_9

Siro, C., Aliannejadi, M., & De Rijke, M. (2023). Understanding and Predicting User Satisfaction with Conversational Recommender Systems. *ACM Transactions on Office Information Systems*, *42*(2), 1–37. https://doi.org/10.1145/3624989

Sondakh, D. E. S., Taju, S. W., Yuune, J. K. T., Kaminang, A. F., & Wagey, S. G. (2024). Research Project Topic Recommender System using Generative Language Model. *CogITo Smart Journal*, *10*(1), 233–245. https://doi.org/10.31154/cogito.v10i1.678.654-666

Sordoni, A., Galley, M., Auli, M., Brockett, C., Ji, Y., Mitchell, M., Nie, J., Gao, J., & Dolan, B. (2015). A Neural Network approach to Context-Sensitive Generation of Conversational Responses. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, (pp. 196–205). ACL. https://doi.org/10.3115/v1/n15-1020

Thiprak, S., & Kurutach, W. (2015). Ubiquitous computing technologies and context aware recommender systems for ubiquitous learning. In *2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON).* IEEE. https://doi.org/10.1109/ECTICon.2015.7206939

Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., & Duval, E. (2012). Context-Aware Recommender Systems for Learning: A survey and future challenges. *IEEE Transactions on Learning Technologies*, *5*(4), 318–335. https://doi.org/10.1109/tlt.2012.11

Wang, S., & Wu, C. (2011). Application of context-aware and personalized recommendation to implement an adaptive ubiquitous learning system. *Expert Systems With Applications*, *38*(9), 10831–10838. https://doi.org/10.1016/j.eswa.2011.02.083

Waswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In *31st Conference on Neural Information Processing Systems (NIPS 2017)*. NIPS. https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Zhai, X. (2022). ChatGPT User Experience: Implications for Education. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4312418

Zhang, Y., Jin, Y., Chen, L., & Yang, T. (2024). Navigating User experience of ChatGPT-based Conversational Recommender Systems: The effects of Prompt Guidance and Recommendation Domain. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2405.13560

Acta Informatica Pragensia is published by the Prague University of Economics and Business, Czech Republic | eISSN: 1805-4951