Smart Start: A Critical Analysis of Chatbots as Tools to Support Study Orientation

Lea Cohausz [1], Thilo I. Dieing [1], Belinda Merkle [2], Marc Scheffler [1], and Stefan Janke [2]

Abstract: In higher education, elevated dropout rates pose a significant problem, leading to vast societal and personal costs and disadvantages. In Germany, in particular, dropout is closely connected to a misfit between students' expectations and their choice of study program. As a countermeasure, a chatbot is proposed that targets prospective students and is supposed to help them with their study orientation by providing them with accurate and personalized information. In this study, we first use a survey including the Technology Acceptance Model to test whether the target group would even accept such a chatbot and whether there are differences in the acceptance for different demographic groups. Our results show high acceptance and no differences among demographic groups, although we propose that differences among demographic groups should be scrutinized in future research. Subsequently, we develop a chatbot architecture using Large Language Models and Retrieval Augmented Generation that enables accurate and personalized answers while simultaneously ensuring appropriate responses to inappropriate user prompts. Our first evaluation of the chatbot shows promising results. The chatbot will become part of a study orientation platform in one of Germany's largest federal states.

Keywords: chatbot, LLM, dropout, service-oriented chatbot, study choice advisor, TAM

1 Introduction

Dropout rates in higher education institutions are high in many countries, particularly in systems that require students to choose a major immediately, such as in Germany or Switzerland. In Germany, 28 % of Bachelor students drop out [He22]. These dropouts incur both societal costs (wasted financial resources) and personal costs for the individuals involved (frustration, time loss)[3] [Fa18].

Although multiple factors contribute to dropping out, a German study by Blüthmann et al. found that the main reasons are a mismatch between study program content and personal interests (35.5 %) or misjudging the program's demands (25.3 %) [BLT12]. Other research

1 University of Mannheim, Chair of Artificial Intelligence, 68159 Mannheim, Germany,
lea.cohausz@uni-mannheim.de, https://orcid.org/0000-0002-6164-3988;
thilo.dieing@uni-mannheim.de, https://orcid.org/0009-0000-2256-0795;
marcscheffler@outlook.com, https://orcid.org/0009-0004-1249-3487

2 University of Mannheim, Chair of Educational Psychology, 68159 Mannheim, Germany,
belinda.merkle@uni-mannheim.de, https://orcid.org/000-0003-1799-8850;
stefan.janke@uni-mannheim.de, https://orcid.org/0000-0002-1916-0069

3 The cost of providing one student with higher education in Germany is around 10,500€ annually [Br22]. With 481,962 first-year Bachelor students in 2023 [Bu24] and an average dropout time of 3.8 semesters, this amounts to 2.69 billion euros in taxpayer money and about 256,403 lost study years per cohort [He17].

doi:10.18420/delfi2025_11

also shows that students with a good match between their interests and study content are more satisfied [MMD24]. Therefore, reducing the mismatch by providing relevant information to prospective students could help lower dropout rates. However, the overwhelming amount of information available—ranging from higher education institutions to online self-assessments (OSAs), financing information, and career fairs—can be difficult to navigate. First-generation academics or students from migrant backgrounds may struggle more due to a lack of social and cultural capital, as evidenced by higher dropout rates in these groups [MM22; MS13]. These students may also be reluctant to approach study advisors.

To address this issue, we aim to develop a study advisor tool that will provide answers to key questions (e. g., about programs, admissions, financial aspects, and OSAs), offering additional guidance for school graduates. This tool needs to: (1) be accepted by prospective students, meaning it is perceived as useful, easy to use, and accessible; (2) be equally accepted across demographic groups; (3) answer a wide range of student questions accurately, with links to external materials where needed; and (4) identify and avoid responding to inappropriate questions, such as those that are discriminatory or sensitive.

We propose using Large Language Models (LLMs) to create a chatbot capable of offering personalized, relevant information with a low threshold for use. LLM-based chatbots can answer a variety of questions without needing predefined responses, and with Retrieval Augmented Generation (RAG), they can provide precise answers tailored to specific queries [Fa24]. However, there are challenges when using LLMs in critical information systems. First, it is unclear whether users will accept the guidance of LLM-based chatbots. Second, there is uncertainty about whether acceptance varies across demographic groups, which is important for ensuring the tool's widespread use. Third, LLMs sometimes provide incorrect answers, a phenomenon known as "hallucination" [Fa24]. Fourth, they tend to respond to sensitive or inappropriate questions [Ba21]. In this paper, we aim to address these challenges:

- We investigate prospective students' acceptance of a study advisor chatbot using a survey based on the Technology Acceptance Model (TAM) in Section 3. Our findings suggest high acceptance [BLT12].
- We explore whether acceptance varies across demographic groups in Section 3.2.2. While our data shows no significant differences, we note that further research is needed.
- We justify the need for a reliable study advisor chatbot that provides sensible answers with minimal hallucination in Section 4. Our evaluation shows positive results.
- We describe a filter in our chatbot design that detects inappropriate questions, ensuring it avoids answering them. Section 4.2.2 demonstrates its effectiveness.

This academic investigation is crucial as the resulting study-advisor chatbot will be integrated into a study orientation platform aimed at helping prospective students navigate the complex decision-making process between higher education institutions and the myriad of available study programs.

2 Background

We first discuss existing literature concerning chatbots in general, chatbots in education, and chatbots' acceptance to subsequently derive our research objective.

2.1 Characteristics and Use of Chatbots

Chatbots have become widely used across various fields, from management to IT support [AM20]. They can be categorized based on four key characteristics: openness, service, goal, and functionality [AM20]. Openness refers to whether a chatbot addresses a broad range of topics (open) or a specific one (closed). Service indicates whether the chatbot offers general or personalized interactions. Goal defines whether the chatbot's purpose is to inform, engage in conversation, or perform tasks. Functionality relates to how the chatbot generates responses, either through rule-based methods or generative Machine Learning, with or without additional information (RAG).

With the rise of advanced LLMs in recent years, most modern chatbots are classified as open or semi-open. These chatbots typically offer personalization, can inform, converse, and solve tasks, and rely on generative Machine Learning along with Retrieval Augmented Generation (RAG), which incorporates domain-specific information to enhance the chatbot's output [Fa24].

2.2 Chatbots in Education

A 2023 survey by Kuahil et al. found that 88.88 % of chatbots used in education are rule-based, relying on predefined question-answer pairs [Ku23]. This limits their functionality, as closed chatbots are suited only for narrow contexts. In contrast, a study-advisor chatbot needs to address a wide range of specific and broad questions. While the number of open chatbots is expected to grow quickly, this finding highlights the current limitations of educational chatbots.

Educational chatbots can also be classified by purpose, either teaching- or service-oriented [OA21; PDP20]. Teaching-oriented chatbots support students and educators in learning tasks, such as teaching coding [DTZ20; Wi20], engineering [Me20], or grading [Ba24]. Service-oriented chatbots, on the other hand, assist students and administrative staff with daily tasks, like helping with curriculum or study decisions [OA21]. A study-advisor chatbot is a clear example of a service-oriented chatbot. While service-oriented chatbots do exist in education [Di18; FSW23], they are less common than teaching-oriented ones [OA21]. The adoption of LLM-based solutions has not yet fully reached this area. As a result, most service-oriented chatbots remain rule-based and are limited to closed knowledge domains [Di18; FSW23]. LLMs supported by RAG offer significant benefits, especially for study-advisor chatbots, as they need to answer a variety of questions on diverse topics and levels of specificity, such as “How much money do I need per month while studying?” and “How much statistics is required in program X at university Y?”

2.3 Chatbots and their Acceptance

Chatbots can positively impact various sectors, including businesses [AMA23; KK20], healthcare [GSK21; Mi21], and education [DSC24; RRS17]. However, for chatbots to realize their full potential, users must engage with them, trust their guidance, and incorporate the information into their decision-making. In the case of study-advisor chatbots, educational effectiveness depends on users’ trust in the information and their integration of it into their study decisions. The Technology Acceptance Model (TAM) is useful for evaluating user acceptance [Da89]. According to TAM, acceptance of a system is based on the perceived benefits it offers, divided into two factors: Perceived Usefulness (PU), which measures how much the system improves the current state, and Perceived Ease of Use (PEU), which reflects the effort required to use the system. PU indicates the system’s utility, so a study-advisor chatbot would be considered useful if it aids the study decision process. PEU describes the ease of navigating the system, and systems that are easier to use tend to be more effective. High PU and PEU lead to a positive Attitude toward using the system (ATT), which influences Behavioral Intentions (BI) and, eventually, actual system usage. However, actual use can only be measured once the system is developed. As developing such systems is costly, it is crucial to first assess general acceptance before proceeding with development. Figure 1 illustrates the causal relationships in the TAM model [Da89], and the model’s components are typically evaluated through user surveys.

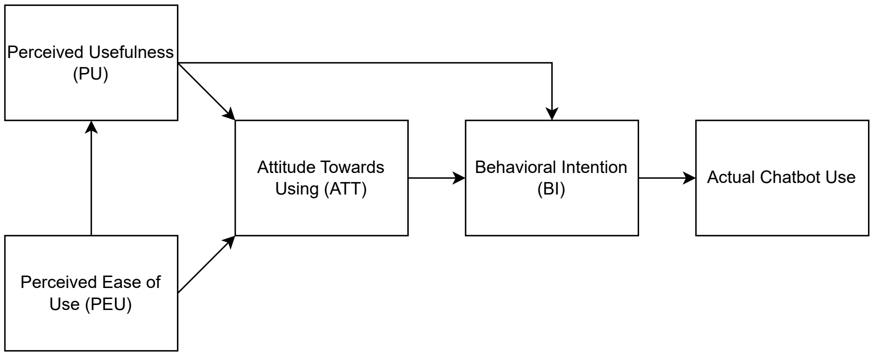

Fig. 1: Technology Acceptance Model based on Davis [Da89].

TAM has been widely used to evaluate chatbot acceptance, including in sectors like banking [AVS21] and tourism [SA24]. In education, studies applying the TAM framework highlight the significance of PU as a key factor in students' [LCC23] and teachers' [CCM23] willingness to use educational chatbots [BKS21]. Further research has identified potential gender differences in chatbot acceptance [Yi23], with male students generally more willing to use chatbots and expressing fewer concerns [SOM24], while other studies find no significant effects of gender or age [Ra22].

2.4 Research Objectives

Considering existing literature and guided by our requirements formulated in Section 1, we formulate the following research questions: **RQ1:** Do student aspirants in the study orientation phase accept a study-advisor chatbot according to the concepts covered by the TAM? **RQ2:** Are there group differences for any of the concepts covered by the TAM for different demographic groups (gender, first-generation Germans, parental education, age)? **RQ3:** If **RQ1** is answered affirmatively: Can we construct a chatbot using open-source LLMs that can answer a wide range of questions correctly and precisely? **RQ4:** If **RQ1** is answered affirmatively: Is it possible to filter inappropriate questions reliably?

As our later two research questions hinge on a positive answer to **RQ1**, we will first describe the methodology for answering the first two questions and then describe the respective results. Afterward, we will discuss the development and evaluation of the chatbot.

3 Chatbot Acceptance

3.1 Method

To address our first research goal—measuring the user acceptance of a study-advisor chatbot—we conducted an online survey to gather opinions from potential users. Our target group included individuals who are or have been in the process of choosing a study program. This group consisted of young adults still attending school and those who have recently decided on their study program, including those in their first semesters. We focused on the latter group, as they have recently navigated the decision process and can provide insights on what would have been helpful to them.

The survey included various question categories based on the TAM model, incorporating factors that influence system usage, and adapted from Ahn et al. [ARH07] to suit a chatbot and study orientation context. Perceived Usefulness (PU) was assessed with six statements, such as "I anticipate that the chatbot is helpful for my study orientation." Perceived Ease of Use (PEU) was evaluated through four statements, e. g., "I anticipate that the chatbot is easy to use." User Attitude (ATT) was captured with one statement, and Behavioral Intention (BI) with two. All items used a 7-point scale, ranging from strong disagreement to strong agreement. We also collected demographic data, including gender, age, education, parental education, and immigration background. Additionally, we asked about participants' use of ChatGPT, as we hypothesized that frequent use of chatbots could influence acceptance of the study-advisor chatbot, making users more likely to view chatbots as useful and be open to using others.

We used a convenience snowball sampling method to reach the target group, collecting data from August 18th to September 9th, 2024. The final sample included 80 participants, with 66 fitting our target group and 65 passing an attention test, which ensured that participants answered thoughtfully [MK16]. Of the 65 valid responses, 55 % identified as female, 43 % as male, and the rest as non-binary. The average age was 19.3 years (SD = 2.3), with ages ranging from 16 to 24. Among the participants, 65 % had recently completed their study orientation and started a study program or vocational training, while the others were in or near their study orientation phase. A notable observation from the demographic data is the relatively high level of parental education and low percentage of first-generation Germans, reflecting our focus on young people having achieved or in the processes of achieving a general university entrance qualification (Abitur) [MM22]. At this point in their lives, the social and cultural capital associated with parental education and migration background has already influenced their educational attainment [MM22].

Next, we assessed the reliability of the constructs using Cronbach's alpha and Confirmatory Factor Analysis (CFA) [ACS15; Br15; Cr51]. The results indicated that the items for each construct were reliably measuring their respective constructs. All constructs had Cronbach's alpha values above 0.8, indicating strong reliability. The model fit was also moderate to good: χ^2, $p < 0.000$, RMSEA = 0.014, SRMR = 0.084, CFI = 0.888. We found that some items from PEU and PU also loaded on the other construct, but these cross-loadings were still in an acceptable range, and some covariance between the constructs was expected.

To examine the impact of demographics and ChatGPT usage on the constructs and their interrelationships, we used path analysis. Path analysis extends multiple regression to model direct and indirect relationships between variables in a hypothesized model [Ro12]. It uses simultaneous linear equations to analyze variable dependencies and provides a visual representation through a path diagram. In our analysis, we specified four structural equations, each with one latent construct as the independent variable. The dependent variables (DVs) included demographic factors and other latent constructs from the TAM model. For example, the equation for BI is: $BI = ATT + PU + Age + Gender + FirstGenerationGerman + ParentWithUniversityDegree + ChatGPT$. This approach allowed us to assess how each latent construct was influenced by both demographic variables and other constructs, accounting for interdependencies. We report fully standardized coefficients for easier comparison.

Variable	Mean/ Mode	Minimum / % Smallest Group	Maximum / % Largest Group	N unique values
Age	19.3	16	24	9
Gender	Female	Non-Binary, 2 %	Female, 55 %	3
First Generation Germans	No	Yes, 18%	No, 82 %	2
Any Parent with University Degree	Yes	No, 35 %	Yes, 65 %	2
ChatGPT	Yes	No, 45 %	Yes, 55 %	2

Tab. 1: The table provides key statistics for variables included in the survey.

3.2 Results

3.2.1 General Chatbot Acceptance

Figure 2 shows the average agreement values for each TAM construct, along with their confidence intervals. The light blue areas represent the response distribution. A value of four reflects a neutral stance, with higher values indicating positive and lower values negative stances. All constructs exhibit similar and very positive average values, ranging from 5.60 (ATT) to 5.84 (PEU). Only a small percentage of participants expressed negative views of the chatbot (7.69 % for ATT and BI, 4.63 % for PEU, and 6.15 % for PU).

Regarding **RQ1**, these results demonstrate high acceptance of the study advisor chatbot according to the TAM model, suggesting favorable prospects for its development and use.

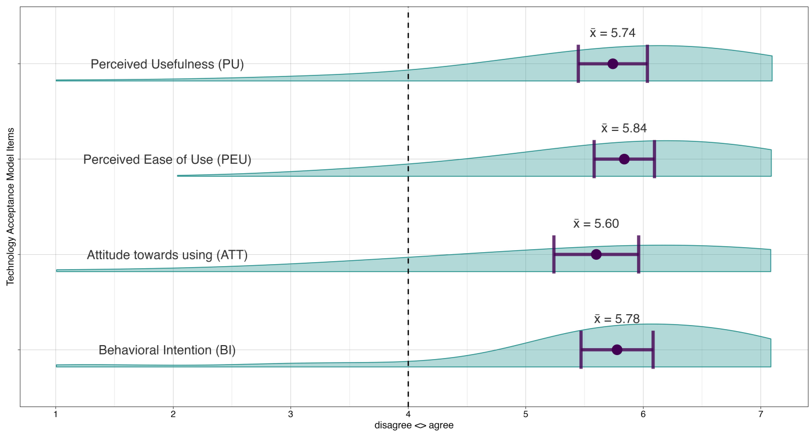

Fig. 2: Acceptance of a study orientation chatbot using the Technology Acceptance Model (TAM) with values ranging from 1 (completely disagree) to 7 (completely agree).

3.2.2 Demographics and Chatbot Acceptance

Table 2 presents the results of the regressions in the path analysis for each TAM construct. Overall, the relationships among constructs align with expectations, though PEU is not significant for ATT, and PU is not significant for BI. This absence of statistical significance may be due to the fact that the other constructs capture much of the variance relevant for the respective dependent variables and that, given the small sample size, smaller remaining variances cannot be utilized. All other relationships in the TAM are significant and strong. For example, a one standard deviation increase in ATT leads to a 0.78 standard deviation increase in BI, on average. Other variables have smaller coefficients in each model.

Age is significant for BI, with a relatively large coefficient for both BI and PEU compared to other demographic variables. The positive coefficient suggests that older individuals may show more agreement, which is surprising as we expected younger individuals to have a higher intention to use the chatbot. This finding may be influenced by the small sample size and limited age range. Other demographic variables are not significant. Although not statistically significant, *First Generation German* shows a relatively large coefficient for PEU. Given the small number of participants in this group, this variable should not be disregarded entirely. Similarly, frequent ChatGPT use has a large coefficient for each construct, although it lacks statistical significance. The coefficient for PEU, PU, and ATT is among the largest, suggesting that a more refined measure of ChatGPT use in a larger sample could reveal interesting patterns.

In conclusion, demographic variables do not show significant associations with the chatbot's acceptance. Only age was significant for BI, while other factors like gender, first-generation status, and parental education were not. However, caution is needed when drawing conclusions. Thus, **RQ2** can be answered by noting that there are few indications of demographic differences in our sample, though further investigation with larger samples is needed. Additionally, we only measured user expectations and behavioral intention, not actual chatbot usage. Future research should assess a larger sample post-launch to connect expectations with real interaction patterns.

4 Chatbot Development and Evaluation

Given the positive results for RQ1, which show that prospective students highly accept study advisor chatbots, we now focus on the development and evaluation of a chatbot designed to answer a broad range of questions accurately (RQ3) and filter inappropriate questions effectively (RQ4).

Variables	PEU	PU	ATT	BI
ATT	-	-	-	**0.78 (0.000)**
PU	-	-	**0.71 (0.000)**	0.05 (0.604)
PEU	-	**0.50 (0.000)**	-0.06 (0.578)	-
Age	-0.19 (0.112)	-0.04 (0.724)	0.02 (0.856)	**0.16 (0.026)**
Gender	0.06 (0.643)	0.05 (0.680)	-0.08 (0.399)	0.03 (0.692)
First Generation German	0.15 (0.209)	-0.08 (0.440)	0.10 (0.236)	0.07 (0.323)
Any Parent with University Degree	-0.03 (0.806)	0.02 (0.840)	0.12 (0.193)	-0.09 (0.170)
ChatGPT	0.19 (0.111)	0.24 (0.389)	0.12 (0.181)	0.07 (0.342)

Tab. 2: The table displays the regression outputs of the path analysis for each of the dependent variables PEU, PU, ATT, and BI (columns). Values significant on the 5 %-level are highlighted in bold. The p-values are added in the brackets. The coefficients are standardized.

4.1 Method

4.1.1 Chatbot Architecture

We built a generative open-domain chatbot using open-source Large Language Models (LLMs) combined with Retrieval Augmented Generation (RAG). The user's query is first embedded and compared to an external data source, also embedded.

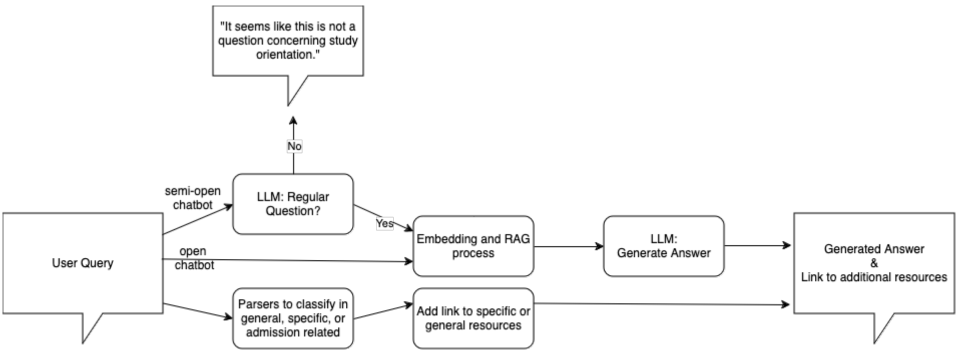

Fig. 3: The architecture of open and semi-open chatbots for study program orientation utilizing RAG and LLM.

This comparison generates a set of best-matching data entries, which are then provided to the LLM to answer the user's question. We tested two RAG match sizes, $N = 30$ and $N = 60$, to handle broader user queries. The more effective size will be used for a semi-open chatbot. This architecture makes the chatbot open, interpersonal, informative/chat-based, and generative. In addition to the open-domain chatbot, we propose a semi-open chatbot with a more limited scope to avoid misuse and prevent the generation of offensive content, as discussed in Section 1 [Ba21]. The semi-open chatbot focuses solely on assisting student aspirants with their decision process. To prevent inappropriate user inquiries, we filter out rude, discriminatory, or sensitive queries before passing them to the LLM. This is illustrated in Figure 3. Unlike the open-domain chatbot, the semi-open bot first evaluates the user's question with an LLM to determine if it is valid. Inappropriate prompts trigger a predefined response, following the topic-avoidance strategy [Xu21]. Combining RAG and LLM ensures factual accuracy while leveraging natural language capabilities. Using an external database improves control over information, reducing hallucinations and avoiding costly pre-training.

For building the chatbots for a German application, we used the Cross English & German RoBERTa Sentence embedding model [Ma20] and Mistralai's Mixtral 8x7B [Ji24]. These models were selected for their strong performance in German-language applications [DSC24]. Open-source models also enable more reproducible results. We sourced data from the Federal Employment Agency's website[4], which includes program names, descriptions, and university details, all embedded for retrieval. A pipeline was added to detect universities and programs in user queries, providing reliable links to information tools like OSAs, ensuring accurate results without hallucinations.

4.1.2 Chatbot Evaluation Standards

Evaluating LLM-driven chatbots remains an open research challenge with no perfect gold standard. The most reliable method is manual grading by coders [DSC24], though it is time-consuming. We assessed open and semi-open chatbots using 70 study orientation questions. These questions consisted of regular prompts (e. g., "Where can I study program X?"), explorative prompts, unusual prompts, including Reddit questions and critical prompts to test the developed filter layer, and prompts provided by real study advisors. Three independent coders then graded responses on a 3-point scale: 1 (correct), 0 (sufficient), and -1 (wrong).

4.2 Results

4.2.1 Overall Results

The coders presented overall agreement in their grades with a Krippendorff's alpha [HK07; Kr18] of 0.70. To evaluate the three tested chatbots, the individual grades were averaged. Figure 4 displays the performances of the bots grouped by the different prompt types. First, we examine the results for the two sizes of RAG matches in the open-domain chatbots. We can see that there are high levels of uncertainty in the answer quality if only 30 matches are given to the LLM. In contrast, the chatbot with 60 matches is able to generate more stable

[4] https://web.arbeitsagentur.de/berufenet/studienfelder

and correct answers to the various user prompts. Based on this, we chose to implement the RAG with 60 matches in the semi-open chatbot. Comparing the open and semi-open bots revealed no significant difference in their ability to answer regular, explorative, or study advisor prompts correctly.

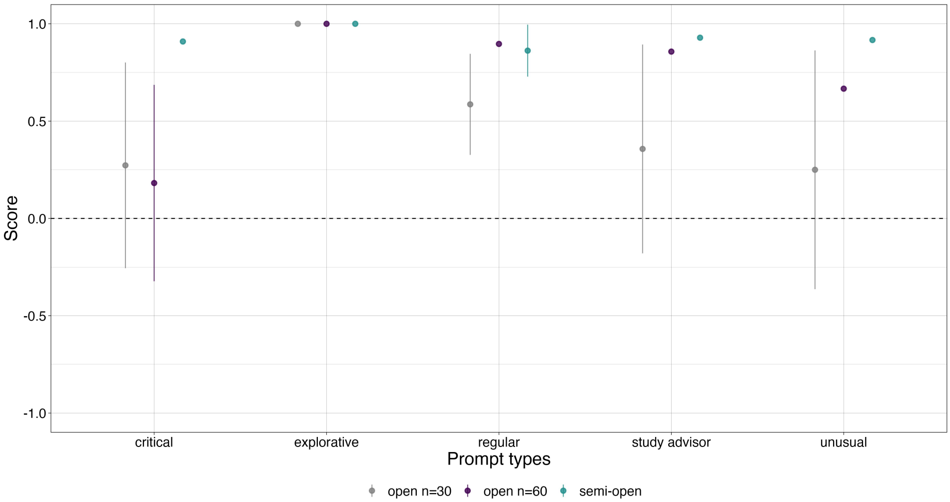

Fig. 4: Performance of both open chatbot versions with 30 and 60 RAG matches and the semi-open chatbot (RAG matches = 60) based on 70 example prompts.

4.2.2 Filtering Inappropriate Questions

Despite strong overall performance, the open-domain chatbot only achieves an average grade of 0.18 for critical prompts, highlighting the need for a filter to manage problematic queries. By applying our filter layer, the semi-open chatbot significantly improves, reaching an average grade of 0.91. These results support the effectiveness of filtering problematic prompts to address potential LLM biases, answering RQ4. Overall, the chatbot evaluation confirms that, as stated in RQ3, using RAG and LLMs is a valid and reliable approach for building a chatbot for study program orientation.

5 Conclusion, Limitations, and Outlook

This paper addresses high dropout rates linked to mismatched student expectations and chosen study programs by proposing a chatbot to support prospective students during their study orientation. The chatbot provides personalized, accurate information to help guide their decisions. Our survey results show high acceptance of the chatbot, with no significant differences across demographic groups. We also outline the rationale for a reliable study-advisor chatbot that minimizes errors, handles inappropriate questions, and delivers sensible answers on study choice.

While the results are promising, this paper serves as a starting point for further long-term evaluation and implementation. Given the positive feedback, we are confident in providing the chatbot to the target group. However, real-world interactions may reveal unforeseen issues that need addressing. These interactions will also shed light on the chatbot's usefulness and acceptance as an additional resource for prospective students. We plan to expand the survey to include more chatbot users and non-users for more accurate insights into demographic differences in usage and acceptance. A limitation of our current study is the small sample size, which restricts our ability to examine demographic factors in depth, such as the effect of being a first-generation German student. A closer look at the parents' countries of origin would also be valuable, as parents from European countries may have more insight into the study orientation process than those from countries with different higher education systems. Further analysis of parental education and prior chatbot experience is also warranted. Future research should explore whether chatbot users find the chatbot helpful in decision-making and, in the long term, whether these users experience lower dropout rates compared to non-users, which is the ultimate goal of the tool. Finally, our evaluation of the developed chatbot demonstrates the potential of this approach, but there is room for improvement. Using the LLM in the filtering layer may introduce biases, so we suggest developing tailor-made classifiers based on a larger dataset of user queries to improve results and reliability.

References

[ACS15] Armentano, M. G.; Christensen, I.; Schiaffino, S.: Applying the technology acceptance model to evaluation of recommender systems. Polibits (51), pp. 73–79, 2015.

[AM20] Adamopoulou, E.; Moussiades, L.: An overview of chatbot technology. In: IFIP international conference on artificial intelligence applications and innovations. Springer, pp. 373–383, 2020.

[AMA23] Azmi, M.; Mansour, A.; Azmi, C.: A Context-Aware Empowering Business with AI: Case of Chatbots in Business Intelligence Systems. Procedia Computer Science 224, pp. 479–484, 2023.

[ARH07] Ahn, T.; Ryu, S.; Han, I.: The impact of Web quality and playfulness on user acceptance of online retailing. Information & management 44 (3), pp. 263–275, 2007.

[AVS21] Alt, M.-A.; Vizeli, I.; Săplăcan, Z.: Banking with a chatbot–A study on technology acceptance. Studia Universitatis Babes-Bolyai Oeconomica 66 (1), pp. 13–35, 2021.

[Ba21] Bang, Y. et al.: Assessing Political Prudence of Open-domain Chatbots, 2021, arXiv: 2106.06157 [cs.CL], https://arxiv.org/abs/2106.06157.

[Ba24] Baral, S. et al.: Automated Assessment in Math Education: A Comparative Analysis of LLMs for Open-Ended Responses. In: Proceedings of the 17th International Conference on Educational Data Mining. Pp. 732–737, 2024.

[BKS21] Brachten, F.; Kissmer, T.; Stieglitz, S.: The acceptance of chatbots in an enterprise context–A survey study. International Journal of Information Management 60, p. 102375, 2021.

[BLT12] Blüthmann, I.; Lepa, S.; Thiel, F.: Überfordert, enttäuscht, verwählt oder strategisch? Eine Typologie vorzeitig exmatrikulierter Bachelorstudierender. Zeitschrift für Pädagogik 58 (1), pp. 89–108, 2012.

[Br15] Brown, T. A.: Confirmatory factor analysis for applied research. Guilford publications, 2015.

[Br22] Brugger, P. et al.: Bildungsfinanzbericht 2015. 2022.

[Bu24] Bundesamt, S.: Anzahl der Studienanfänger:innen im ersten Hochschulsemester in Deutschland in den Studienjahren 1995/96 bis 2023/24. Statista, 2024, https://de.statista.com/statistik/daten/studie/4907/umfrage/studienanfaenger-in-deutschland-seit-1995/.

[CCM23] Chocarro, R.; Cortiñas, M.; Marcos-Matás, G.: Teachers' attitudes towards chatbots in education: a technology acceptance model approach considering the effect of social language, bot proactiveness, and users' characteristics. Educational Studies 49 (2), pp. 295–313, 2023.

[Cr51] Cronbach, L. J.: Coefficient alpha and the internal structure of tests. psychometrika 16 (3), pp. 297–334, 1951.

[Da89] Davis, F. D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, pp. 319–340, 1989.

[Di18] Dibitonto, M. et al.: Chatbot in a Campus Environment: Design of LiSA, a Virtual Assistant to Help Students in Their University Life. In (Kurosu, M., ed.): Human-Computer Interaction. Interaction Technologies. Springer International Publishing, Cham, pp. 103–116, 2018.

[DSC24] Dieing, T. I.; Scheffler, M.; Cohausz, L.: Enhancing Chatbot-Assisted Study Program Orientation. In: Proceedings of DELFI Workshops 2024. Gesellschaft für Informatik eV, pp. 10–18420, 2024.

[DTZ20] Daud, S. H. M.; Teo, N. H. I.; Zain, N. H. M.: Ejava chatbot for learning programming language: Apost-pandemic alternative virtual tutor. International Journal 8 (7), pp. 3290–3298, 2020.

[Fa18] Faas, C. et al.: Socioeconomic success and mental health profiles of young adults who drop out of college. Journal of Youth Studies, 21(5), 669–686, 2018.

[Fa24] Fan, W. et al.: A Survey on RAG Meeting LLMs: Towards Retrieval-Augmented Large Language Models. In: Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Pp. 6491–6501, 2024.

[FSW23] Fersch, M.-L.; Schacht, S.; Woldai, B.: Exploring AI in Education: A Quantitative Study of a Service-Oriented University Chatbot. In: The Paris Conference on Education 2023: Official Conference Proceedings. Pp. 2758–0962, 2023.

[GSK21] Gupta, J.; Singh, V.; Kumar, I.: Florence-a health care chatbot. In: 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). Vol. 1, IEEE, pp. 504–508, 2021.

[He17] Heublein, U. et al.: Zwischen Studienerwartungen und Studienwirklichkeit: Ursachen des Studienabbruchs, beruflicher Verbleib der Studienabbrecherinnen und Studienabbrecher und Entwicklung der Studienabbruchquote an deutschen Hochschulen (Forum Hochschule, 1/2017). Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, 2017.

[He22] Heublein, U.: Die Entwicklung der Studienabbruchquoten in Deutschland. 2022.

[HK07] Hayes, A. F.; Krippendorff, K.: Answering the Call for a Standard Reliability Measure for Coding Data. Communication Methods and Measures 1 (1), pp. 77–89, 2007, DOI: 10.1080/19312450709336664.

[Ji24] Jiang, A. Q. et al.: Mixtral of Experts, 2024, arXiv: 2401.04088 [cs.LG].

[KK20] Kushwaha, A. K.; Kar, A. K.: Language model-driven chatbot for business to address marketing and selection of products. In: Re-imagining Diffusion and Adoption of Information Technology and Systems: A Continuing Conversation: IFIP WG 8.6 International Conference on Transfer and Diffusion of IT, TDIT 2020, Tiruchirappalli, India, December 18–19, 2020, Proceedings, Part I. Springer, pp. 16–28, 2020.

[Kr18] Krippendorff, K.: Content analysis: An introduction to its methodology. Sage publications, 2018.

[Ku23] Kuhail, M. A. et al.: Interacting with educational chatbots: A systematic review. Education and Information Technologies 28 (1), pp. 973–1018, 2023.

[LCC23] Lai, C. Y.; Cheung, K. Y.; Chan, C. S.: Exploring the role of intrinsic motivation in ChatGPT adoption to support active learning: An extension of the technology acceptance model. Computers and Education: Artificial Intelligence 5, p. 100178, 2023.

[Ma20] May, P.: Cross English & German RoBERTa for Sentence Embeddings. 2020, https://huggingface.co/T-Systems-onsite/cross-en-de-roberta-sentence-transformer.

[Me20] Mendez, S. et al.: Chatbots: A tool to supplement the future faculty mentoring of doctoral engineering students. International Journal of Doctoral Studies 15, 2020.

[Mi21] Mittal, M. et al.: Web-based chatbot for frequently asked queries (FAQ) in hospitals. Journal of Taibah University Medical Sciences 16 (5), pp. 740–746, 2021.

[MK16] Merkle, B.; Kaczmirek, L.: Du kommst hier nicht rein: Türsteherfragen identifizieren nachlässige Teilnehmer in Online-Umfragen. Planung und Analyse: Zeitschrift für Marktforschung und Marketing, 1, 2–5. 2016.

[MM22] Mishra, S.; Müller, L.: Resources, norms, and dropout intentions of migrant students in Germany: The role of social networks and social capital. Studies in Higher Education 47 (8), pp. 1666–1680, 2022.

[MMD24] Merkle, B.; Messerer, L. A. S.; Dickhäuser, O.: Will I be happy in this major? Predicting intrinsic motivation and subjective well-being with prospective students' well-being forecast and interest-major fit forecast. Social Psychology of Education, 27(1), 237–259, 2024.

[MS13] Müller, S.; Schneider, T.: Educational pathways and dropout from higher education in Germany. Longitudinal and Life Course Studies 4 (3), pp. 218–241, 2013.

[OA21] Okonkwo, C. W.; Ade-Ibijola, A.: Chatbots applications in education: A systematic review. Computers and Education: Artificial Intelligence 2, p. 100033, 2021.

[PDP20] Pérez, J. Q.; Daradoumis, T.; Puig, J. M. M.: Rediscovering the use of chatbots in education: A systematic literature review. Computer Applications in Engineering Education 28 (6), pp. 1549–1565, 2020, DOI: 10.1002/cae.22326.

[Ra22] Ragheb, M. A. et al.: Investigating the acceptance of applying chat-bot (Artificial intelligence) technology among higher education students in Egypt. International Journal of Higher Education Management 8 (2), 2022.

[Ro12] Rosseel, Y.: lavaan: An R package for structural equation modeling. Journal of statistical software 48, pp. 1–36, 2012.

[RRS17] Ranoliya, B. R.; Raghuwanshi, N.; Singh, S.: Chatbot for university related FAQs. In: 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE, pp. 1525–1530, 2017.

[SA24] Solomovich, L.; Abraham, V.: Exploring the influence of ChatGPT on tourism behavior using the technology acceptance model. Tourism Review, 2024.

[SOM24] Stöhr, C.; Ou, A. W.; Malmström, H.: Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. Computers and Education: Artificial Intelligence 7, p. 100259, 2024.

[Wi20] Winkler, R. et al.: Sara, the lecturer: Improving learning in online education with a scaffolding-based conversational agent. In: Proceedings of the 2020 CHI conference on human factors in computing systems. Pp. 1–14, 2020.

[Xu21] Xu, J. et al.: Recipes for Safety in Open-domain Chatbots, 2021, arXiv: 2010.07079 [cs.CL].

[Yi23] Yilmaz, H. et al.: Student attitudes towards Chat GPT: A technology acceptance model survey. International Educational Review 1 (1), pp. 57–83, 2023.