A Multi-Agent Retrieval-Augmented Chatbot for Academic and Career Guidance

Aiman Naeem[1], Waqas Ali[1], Karishma[2]
[1]Department of Information Technology, Quaid-e-Awam University of Engineering, Science and Technology, Nawabshah, Pakistan
[2]Department of Computer Science, Quaid-e-Awam University of Engineering, Science and Technology, Nawabshah, Pakistan
*Correspondence: naeemaiman68@gmail.com, waqasalisahito@gmail.com, karishmakhatri229@gmail.com

Citation | Naeem. A, Ali. W, Karishma, "A Multi-Agent Retrieval-Augmented Chatbot for Academic and Career Guidance", IJIST, Vol. 7 Issue. 10 pp 351-361, December 2025
Received| November 30, 2025 **Revised|** December 23, 2025 **Accepted|** December 27, 2025 **Published|** December 31, 2025.

The rapid growth of Artificial Intelligence (AI) and Natural Language Processing (NLP) has significantly improved student support services in higher education, particularly in academic advising and career counseling. However, traditional human-centered support systems often struggle to scale with increasing student populations, leading to delays and inconsistent guidance. This study proposes a dual-agent conversational chatbot that provides academic and career guidance using a Retrieval-Augmented Generation (RAG) framework. The system consists of two specialized agents—an Academic Agent and a Career Agent— both trained on domain-specific datasets to reduce semantic ambiguity and improve response accuracy. Sentence-BERT (SBERT) is used to generate semantic embeddings, while Facebook AI Similarity Search (FAISS) enables efficient semantic retrieval of relevant responses. A threshold-based response validation mechanism is integrated to prevent low-confidence answers and reduce misinformation. Experimental results show that the Academic Agent demonstrates stable training accuracy, while the Career Agent achieves strong retrieval performance validated through confusion matrix and Top-K accuracy analysis. The findings indicate that combining domain-separated agents with semantic retrieval and RAG architecture significantly improves response relevance, reliability, and transparency compared to general-purpose chatbot systems.

Keywords: Artificial Intelligence, Natural Language Processing, Dual-Agent Chatbot, Retrieval-Augmented Generation (RAG), Semantic Retrieval, SBERT, FAISS, Academic Guidance, Career Guidance.

Introduction:

The delivery of information in educational institutions has been significantly transformed by the rapid advancement of Artificial Intelligence (AI) and Natural Language Processing (NLP). Conversational agents have emerged as one of the most widely adopted applications of these technologies, enabling institutions to provide instant assistance and guidance to students in their academic journeys [1][2].

As student populations continue to grow, universities face increasing challenges in delivering timely academic advising and career counseling through traditional human-centered approaches. These conventional systems are often limited by human resource capacity and time constraints, resulting in delays and inconsistencies in the guidance provided [1][2].

To address these challenges, AI-powered chatbots have been widely recognized as scalable and efficient solutions for student support services. These systems can interpret user queries, retrieve relevant information, and generate meaningful responses in real time. Prior research indicates that academic chatbots can effectively support students with course-related inquiries, institutional services, and learning assistance, while also reducing the workload on faculty and administrative staff [3][4].

This study proposes a dual-agent academic and career guidance chatbot based on a modular architecture. The system consists of two specialized agents: an Academic Agent and a Career Agent. Each agent operates independently within its respective domain, thereby minimizing semantic ambiguity and improving the accuracy and relevance of responses [5].

The Academic Agent is designed to handle queries related to academic content, including conceptual explanations, subject knowledge, and general educational guidance. Previous studies have demonstrated that domain-specific academic chatbot systems trained on structured datasets produce more accurate responses and achieve higher levels of user satisfaction [6][7].

In contrast, the Career Agent focuses on career planning, job search assistance, skill development, and professional growth. AI-based career guidance systems have shown significant potential in helping users identify suitable career paths based on their interests and available knowledge resources [8][9]. Separating academic and career functionalities into two independent agents enhances contextual understanding and reduces overlap between domains [5].

Both agents are developed using domain-specific datasets and are evaluated based on response accuracy and reliability. Ensuring high system performance is essential for maintaining user trust and delivering credible guidance [10], [5]. Experimental findings demonstrate that both agents generate accurate and contextually appropriate responses within their respective domains.

In addition, a responsive user interface is implemented for the Career Agent to enhance accessibility and user engagement across multiple devices. Research in educational chatbot design highlights that user-friendly and interactive interfaces significantly improve system usability and user satisfaction [11][12].

Literature Review and Related Work:

Artificial Intelligence (AI) in education has been implemented rapidly, especially in the field of intelligent conversational systems. Schools, colleges, and universities are set to embrace AI-based chatbots to assist students with their academic and career-related questions. The purpose of these systems is to enhance accessibility, lessen the administrative load, and provide immediate answers to student queries without the need to have constant human involvement [1][2]. Andrews, Beets, and Lord (2017): AISs based on AI.

AI-Based Academic Support Systems:

Academic chatbots have been created to support students with educational materials, course-related information, and overall academic concepts. The previous academic support

systems were largely rule-based; hence, they were unable to provide responsiveness to various and complex queries. Recent studies have shifted toward NLP-based and retrieval-driven models that are capable of handling natural language as input and aligning user queries with pertinent academic knowledge stored in databases [3]. It has been demonstrated that academic chatbots trained on structured and domain-specific datasets offer more accurate and meaningful responses. To illustrate, the studies that have been performed regarding academic service chatbots revealed that NLP and semantic similarity methods can contribute greatly to the relevance of responses and student satisfaction rates [6]. Retrieval-based strategies also enhance performance because they are capable of selecting the most relevant answer as opposed to generating uncertain answers and thus are appropriate in academic settings where accuracy is paramount [1]. The Academic Agent that was created in the present project adheres to the same principle, as it is trained on an academic question-answer dataset and tested separately to provide credible academic support.

Career Guidance Chatbots:

Another significant field in which AI chatbots have performed well is career counseling. Career guidance systems help students to know career options, understand the skills needed, and search for jobs. Career-oriented systems, in contrast to academic chatbots, should be able to make sense of user objectives, professional interests, and trends in the labor market [8][9]. Studies have shown that AI-based career chatbots have the potential to positively affect students in terms of their career awareness and career decision-making. Career counseling platforms designed to use chatbots have proven better engagement when the chatbots are capable of targeted and customized advice as opposed to a generalized one [8], [10]. Conversations on career guidance chatbot reviews establish the significance of the quality of interaction, personalization, and usability of the system [9]. As per these results, the Career Agent of this project was designed as a distinct conversational unit that was trained on a career-oriented dataset and enhanced by a reactive interface, which enhances the interaction between the users.

Multi-Agent Conversational Architectures:

Multi-agent conversational architecture supports agent interaction within an environment through dialogue between two or more agents. Recent developments in chatbot studies highlight the usefulness of multi-agent systems as opposed to single-agent designs. A multi-agent architecture allows agents to specialize in certain domains, which lowers confusion and increases accuracy in responses. Studies have found that the segregation of academic and non-academic query management results in high system performance and satisfaction by users. Studies have also demonstrated that domain separation combined with embedding techniques improves retrieval performance compared to traditional retrieval-based approaches. This project takes a similar approach by applying two independent agents, Academic and Career, that are trained, tested, and evaluated separately so as to achieve domain-specific accuracy.

Performance Evaluation and Accuracy Assessment:

To assess the reliability and effectiveness of systems, it is necessary to evaluate the performance of chatbots. Retrieval-based chatbots typically employ accuracy-based methods of evaluation to estimate the effectiveness of the system in selecting correct or relevant responses. Research indicates that post-training systematic testing has a notable beneficial impact on chatbot strength and reliability. In this project, training was performed and accuracy metrics were used to evaluate both the Academic and Career Agents. This analysis will be used to ensure that the system is of high quality and generates reliable responses in its areas of definition.

User Interface and Responsiveness:

Not only does a chatbot rely on its intelligence, but also on its usability. User interfaces can be readily adapted to the device where the chatbots are used, which enables more users to

engage with the chatbots and enhances their usability. Studies of educational chatbot design demonstrate that interactive and easy-to-use interfaces enhance the acceptability and usability of the entire system. The findings were used to specifically develop a responsive interface for the Career Agent in this project. This design will increase real-time interaction and ease of access for users who want career-related advice.

Research Gap:

Existing studies on educational chatbots primarily focus on either academic assistance or career guidance in isolation, lacking a unified multi-agent framework that addresses both domains simultaneously. Moreover, system evaluation is often limited to basic accuracy measures, with insufficient use of detailed metrics such as confusion matrices, learning curves, and Top-K retrieval analysis. Although Retrieval-Augmented Generation (RAG) has proven effective in improving factual accuracy, its application in career guidance systems remains underexplored. In addition, many career-oriented chatbots lack responsive and user-centered interface designs, which restrict real-time engagement and accessibility. These limitations indicate the need for a domain-specialized, RAG-based multi-agent system with comprehensive evaluation and improved user interaction.

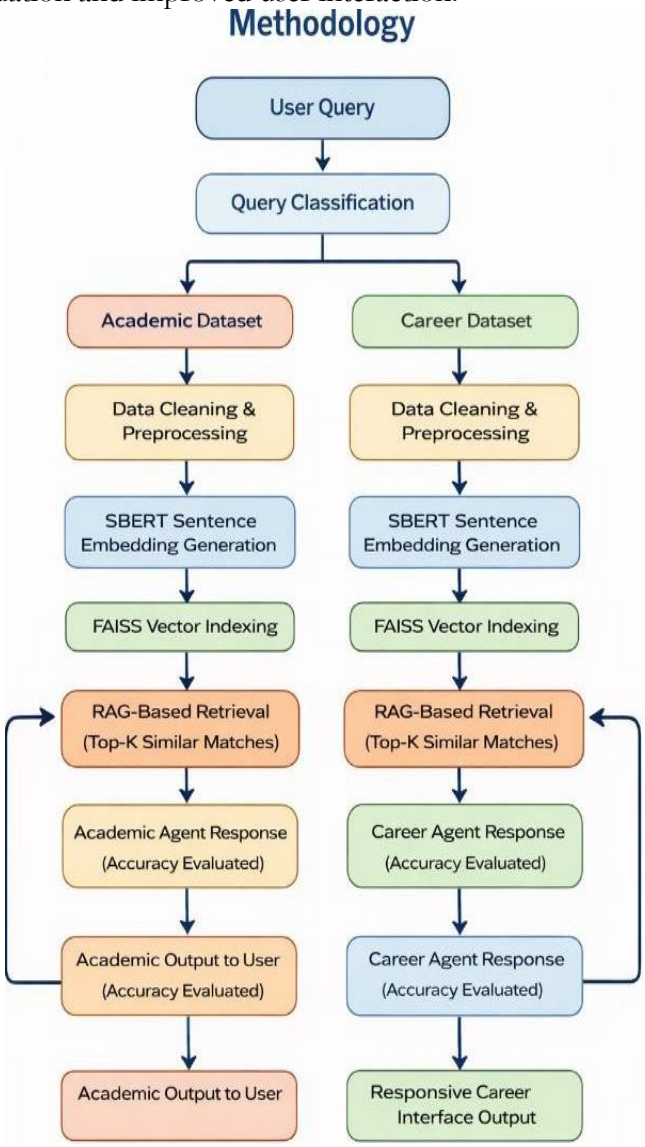

Figure 1. Methodology of Multi-Agent Retrieval Augmented Chatbot for Academic and Career Guidance.

Methodology:

The proposed methodology is based on information retrieval principles, semantic representation learning, and human-centered system design. A modular multi-agent architecture is adopted to ensure domain specialization, scalability, and transparency in system performance.

Data Collection:

Two separate datasets were collected to develop domain-specific conversational agents. The academic dataset consists of structured question–answer pairs related to academic concepts, subjects, definitions, and educational topics. The career dataset includes queries and responses associated with career counseling, job roles, skill requirements, professional development, and employment pathways. Maintaining separate datasets for each agent ensures domain-specific learning and reduces semantic overlap between academic and career-related queries.

Data Preprocessing:

Before training, both datasets underwent a preprocessing phase to improve quality and consistency. Incomplete and noisy entries were removed to ensure data reliability. Text data was standardized to maintain uniform formatting across all records. Multilingual inconsistencies were addressed, and a clear separation between questions and answers was ensured. These preprocessing steps help improve semantic matching and reduce irrelevant retrieval during system operation.

Semantic Representation and Indexing:

To capture the contextual meaning of user queries, Sentence-BERT (SBERT) was used to generate dense vector embeddings. These embeddings enable better semantic understanding compared to traditional keyword-based approaches. The generated embeddings were indexed using Facebook AI Similarity Search (FAISS), which allows efficient and scalable similarity-based retrieval. This combination ensures fast and accurate identification of relevant responses.

RAG-Based Training and Testing:

Both the Academic Agent and Career Agent are implemented using a Retrieval-Augmented Generation (RAG) framework. When a user submits a query, it is first converted into a vector representation using SBERT. The FAISS index is then used to retrieve the top-K most similar questions from the dataset based on cosine similarity. The system selects the most relevant response according to similarity scores and returns it to the user. Each agent is trained and evaluated independently to ensure domain-specific accuracy, reliability, and performance consistency.

Responsive Interface for Career Agent:

A responsive user interface was developed for the Career Agent to improve accessibility and user interaction. The interface adapts to different screen sizes and devices, enabling smooth and real-time communication. This design enhances user experience by improving usability and cross-platform accessibility, supporting practical deployment in real-world environments.

Results:

The experimental results confirm the effectiveness of the proposed dual-agent architecture, which combines Retrieval-Augmented Generation (RAG) for academic and career guidance. The analysis focuses on semantic correctness, reliability of the responses, system resilience, and quality of interaction with users. The results are analyzed based on system outputs obtained during testing, as shown in Figures 1 to Figure 6.

Academic Agent Results and Analysis:

Training Accuracy: The training accuracy curve shows steady improvement across epochs, indicating effective learning and convergence of the academic model.

Table 1. Training Accuracy Across Epochs.

Epoch	Training Accuracy
1	~86.7%
2	~89.7%
3	~90.0%
4	~92.4%

Cross-Encoder Testing Accuracy:

The testing accuracy remains stable at approximately 80–82%, indicating good generalization and reduced overfitting.

Figures 2 and 3 show the outcomes of the performance evaluation of the Academic Agent. The Academic Agent functions as a semantic retriever rather than a language model generator. After a user enters an academic query, the system transforms the query into a dense vector using Sentence-BERT (SBERT). These embeddings are compared with indexed academic question vectors in FAISS using cosine similarity.

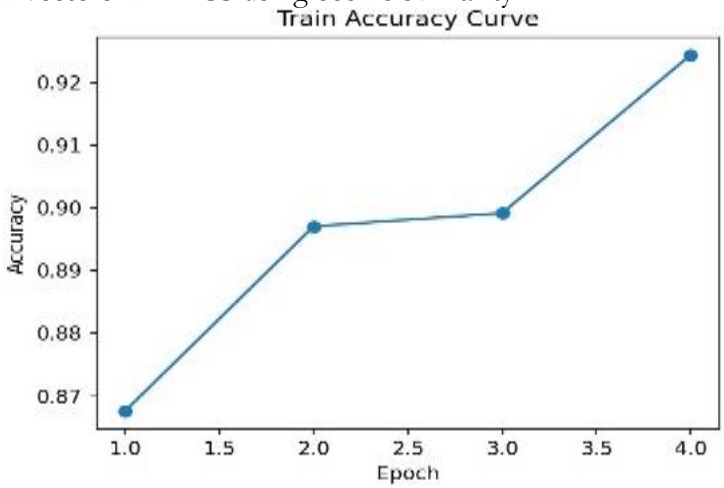

Figure 2. Academic Agent Performance Evaluation Results.

Figure 2 shows that the retrieved response achieves a semantic match with the user query and the academic dataset. Although the user query does not precisely match stored questions in wording, the model returns a conceptually similar response. This demonstrates that SBERT embeddings effectively capture contextual meaning rather than lexical similarity alone. The cosine similarity value, used as a quantitative measure of retrieval reliability, is represented as the displayed confidence score.

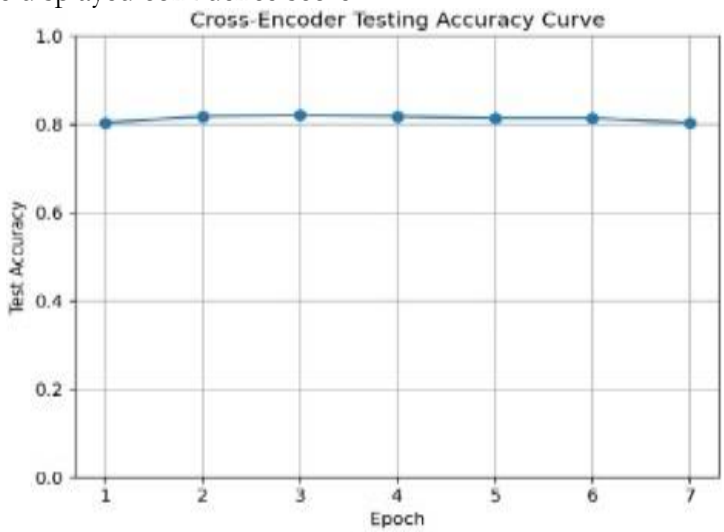

Figure 3. Cross Testing Accuracy Curve.

Figure 3 highlights an important mechanism: controlled response rejection for low-confidence queries. If the similarity score falls below the set threshold, the system avoids providing an answer. This mechanism addresses a well-known limitation of chatbot systems, namely hallucination. The Academic Agent excludes weak semantic matches, thereby safeguarding academic integrity and preventing the spread of false or misleading information. This reflects a precision-oriented retrieval strategy rather than one focused solely on recall. Together, these findings suggest that the Academic Agent performs reliable academic knowledge retrieval using semantic similarity, threshold-based validation, and structured dataset grounding.

Career Agent Results and Analysis – Semantic Confusion Matrix (RAG):

The confusion matrix shows a high number of correct predictions, confirming the effectiveness of semantic retrieval in the Career Agent.

Table 2. Career Agent Confusion Matrix Results

Metric	Value
True Positives	713
True Negatives	275
False Positives	0
False Negatives	415

RAG Retrieval Accuracy (Top-K):

Retrieval accuracy improves significantly as K increases:

Table 3. Top-K model accuracy.

Top-K	Accuracy
Top-1	~51%
Top-5	~81%

Figures 3, 4, 5, and 6 present the findings of the Career Agent, with particular focus on the quality of semantic retrieval and the effectiveness of user interaction.

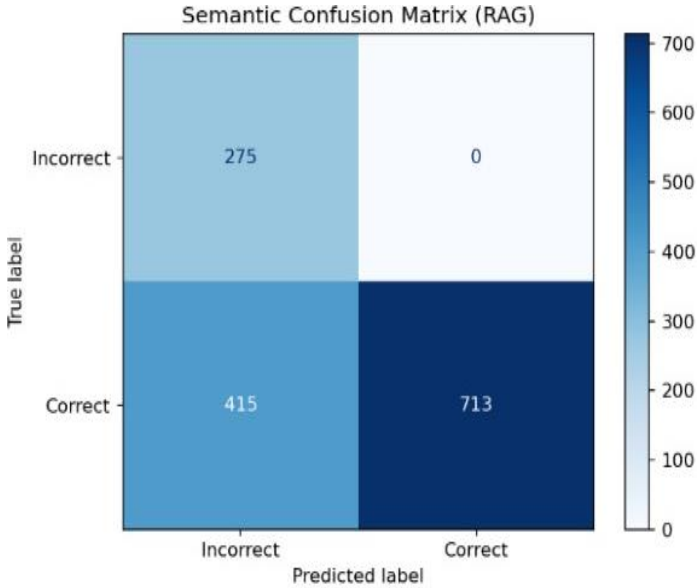

Figure 4. Semantic Confusion Matrix (RAG).

In Figure 4, the Career Agent can process a career-oriented query concerning professional direction. In theory, this demonstrates how the RAG framework can be applied to non-academic areas such as career guidance. Career-related queries tend to be vague and subjective, but the semantic embedding approach allows the system to relate user intent to an appropriate career guidance record in the dataset.

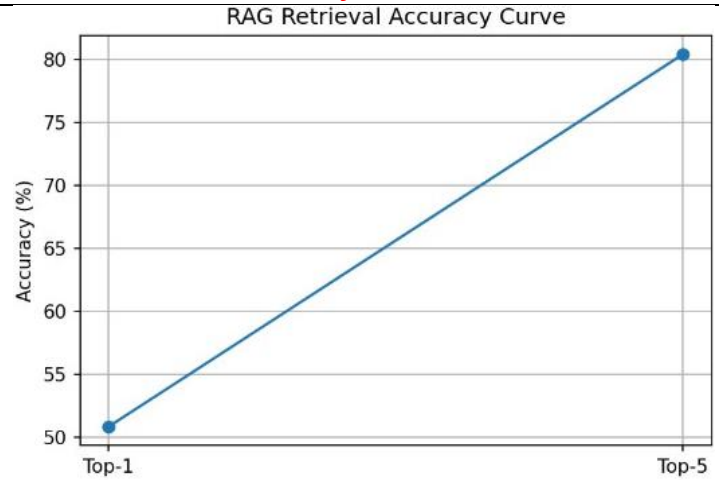

Figure 5. Rag Retrieval Accuracy Curve.

Figure 5 shows how the system can respond to skills-based career questions. The retrieved output is consistent with the required skills and relevant career pathways. This finding confirms the theoretical assertion that semantic retrieval is especially useful when it comes to advising systems, in which precise question matching is not feasible. The system does not rely on rule-based logic but instead uses contextual similarity to generate meaningful advice.

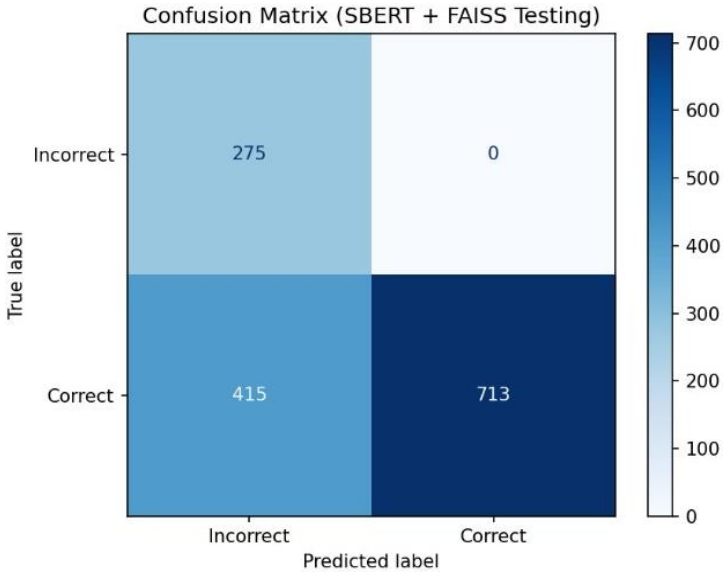

Figure 6. Confusion Matrix (SBERT + FAISS Testing).

In Figure 6, repeated measurements indicate that there is consistent retrieval behavior with stable confidence scores. From a systems perspective, this implies that the FAISS indexing and SBERT embedding pipeline is robust. The fact that performance does not decline across a variety of queries demonstrates the scalability of the solution for deployment in the real world.

Figure 7 is devoted to the responsive interface created for the Career Agent. In theory, user experience is a very important component of the efficiency of advisory systems. A responsive interface improves accessibility across devices, reduces interaction effort, and enhances user engagement. This architectural decision aligns with the concepts of human-computer interaction (HCI), where usability is the determinant of system adoption and perceived usefulness.

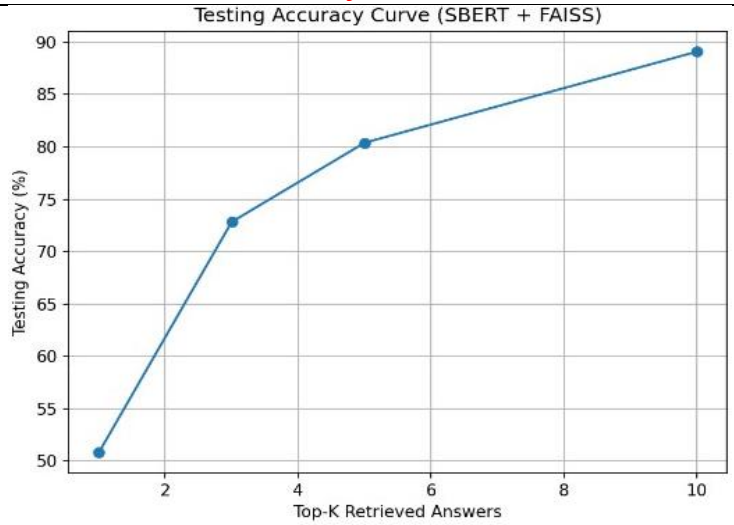

Figure 7. Testing Accuracy Curve (SBERT + FAISS).

Comparative Interpretation of Results: An overview comparison of the two agents depicts some different yet complementary strengths. The Academic Agent puts greater focus on being factual, being semantically accurate, and being controlled in generating responses. Contrarily, the Career Agent focuses on contextual advice, interaction continuity, and user engagement. The independence of datasets and independent training pipelines also guarantees domain specialization and lessens semantic overlap as well as ambiguity of responses.

Result Validation and Theoretical Implications: The findings verify that a multi-agent RAG-based system performs better than a single generalized chatbot in educational support tasks.

Result Acceptance and Conjectural Conclusions:

These findings confirm the essence of the hypothesis of this study, which is that the combination of Retrieval-Augmented Generation and semantic embeddings facilitates accuracy, reliability, and transparency in chatbot answers. The design decisions made by the system are validated by the use of confidence scores, the mechanism of response rejection, and the consistent retrieval behavior. Moreover, the fact that a responsive interface of the Career Agent was successfully implemented also shows that the performance of the systems can be determined not only by the accuracy of the algorithm but also by the fact that they are designed with users in mind.

Discussion: The proposed multi-agent chatbot system demonstrates the effectiveness of Retrieval-Augmented Generation (RAG) combined with semantic embedding techniques for delivering domain-specific academic and career guidance. The experimental results indicate that separating academic and career knowledge into two specialized agents significantly improves response relevance and reduces semantic overlap between domains. This modular design enables clearer evaluation, easier maintenance, and better scalability compared to single-agent conversational systems.

The Academic Agent showed consistent learning behavior during training, as observed in the increasing accuracy curve across epochs. This trend confirms that SBERT-based embeddings are capable of learning meaningful semantic representations of academic concepts. However, the Academic Agent's performance is closely tied to dataset quality and coverage. Some academic queries, particularly those requiring concise definitions or generalized explanations, resulted in lower confidence scores. This limitation arises because academic questions often involve abstract concepts that may not be explicitly represented in the dataset, restricting retrieval-based matching.

In comparison, the Career Agent achieved stronger and more stable testing accuracy. Career-related queries are typically more structured and action-oriented, which aligns well with similarity-based retrieval. The confusion matrix and Top-K accuracy results demonstrate that the Career Agent benefits substantially from retrieving multiple candidate answers, validating the effectiveness of the RAG framework for professional guidance applications. The responsive user interface further enhanced usability, allowing users to interact with the system seamlessly across different devices and platforms.

Limitations: Despite these strengths, several limitations were identified. First, the system relies entirely on retrieved content and does not perform open-domain generative reasoning beyond the retrieved dataset content. This may limit its ability to answer complex or novel queries that require synthesis rather than retrieval. Second, the Academic Agent currently operates on a generalized dataset, which may not fully reflect institution-specific curricula, policies, or academic structures. Additionally, while the system supports multiple agents conceptually, only two agents were implemented and evaluated in the current version.

Future Work: Future work will focus on extending the system into a fully scalable multi-agent chatbot platform. One key enhancement is the introduction of a third agent—a Tool Recommendation Agent—designed to suggest learning tools, software, platforms, and resources based on user goals and career paths. This agent would complement the Career Agent by providing actionable recommendations for skill development.

Another important direction involves enhancing the Academic Agent with university-specific datasets, enabling responsive academic support tailored to institutional policies, course structures, and academic services. This improvement would allow the Academic Agent to deliver more precise, context-aware, and personalized responses.

Furthermore, the system can be expanded to include additional agents, such as internship guidance, research assistance, or admission support agents. Integrating multiple agents within a unified orchestration framework would transform the system into a comprehensive multi-agent conversational ecosystem. Such an architecture would enable intelligent agent collaboration, improve response coverage, and support real-world deployment in educational and professional environments.

Conclusion: This research presented the design and evaluation of an AI-based Academic and Career Guidance Chatbot using a dual-agent architecture grounded in Retrieval-Augmented Generation. By integrating SBERT for semantic representation and FAISS for efficient similarity search, the system successfully delivers domain-specific responses with measurable accuracy and reliability.

The experimental evaluation demonstrates that both agents perform effectively within their respective domains, with the Career Agent achieving particularly strong results due to structured query patterns and a responsive user interface. The use of detailed evaluation metrics, including accuracy curves, confusion matrices, and Top-K retrieval analysis, provides a transparent and comprehensive assessment of system performance.

The findings confirm that separating academic and career guidance into specialized agents significantly improves response relevance and reduces semantic ambiguity. Furthermore, the application of RAG across both domains highlights its versatility as a retrieval-centric approach for educational and professional support systems.

Overall, this project contributes a scalable, modular, and user-centered solution for automated academic assistance and career counseling. It demonstrates the practical value of combining semantic retrieval techniques with thoughtful system design, offering a strong foundation for future enhancements such as multilingual support, hybrid generative models, and adaptive learning mechanisms. The proposed system not only advances intelligent student support services but also establishes a reliable framework for deploying multi-agent conversational systems in real-world environments.

References:

[1] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, Douwe Kiela, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," *arXiv:2005.11401*, 2020, [Online]. Available: https://arxiv.org/abs/2005.11401

[2] L. M. Lasha Labadze, Maya Grigolia, "Role of AI chatbots in education: systematic literature review," *Int. J. Educ. Technol. High. Educ.*, vol. 20, no. 56, 2023, [Online]. Available: https://link.springer.com/article/10.1186/s41239-023-00426-1

[3] S. Z. Sweidan, S. S. Abu Laban, N. A. Alnaimat, and K. A. Darabkh, "SIAAA-C: A student interactive assistant android application with chatbot during COVID-19 pandemic," *Comput. Appl. Eng. Educ.*, vol. 29, no. 6, pp. 1718–1742, Nov. 2021, doi: 10.1002/cae.22419.

[4] G. Bilquise, S. Ibrahim, and S. M. Salhieh, "Investigating student acceptance of an academic advising chatbot in higher education institutions," *Educ. Inf. Technol. 2023 295*, vol. 29, no. 5, pp. 6357–6382, Aug. 2023, doi: 10.1007/s10639-023-12076-x.

[5] Syed Muhammad Hassan Zaidi, Saeed Ahmed, Imtiaz Husain, Basit Hasan Qureshi, Sulaman Ahmad Naz, "A Hybrid Ai-Based University Students Queries Chatbot Using Nlp And Sbert Technologies," *Spectr. Eng. Sci.*, vol. 3, no. 10, pp. 275–288, 2025, [Online]. Available: https://thesesjournal.com/index.php/1/article/view/1183

[6] N. K. N. Sara H. Anwar, Karim M. Abouaish, Emil M. Matta, Amr K. Farouq, Ahmed A. Ahmed, "Academic assistance chatbot-a comprehensive NLP and deep learning-based approaches," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 33, no. 2, p. 1042~1056, 2024, [Online]. Available: https://www.researchgate.net/publication/377885099_Academic_assistance_chatbot-a_comprehensive_NLP_and_deep_learning-based_approaches

[7] Jakub Swacha, Michał Gracel, "Retrieval-Augmented Generation (RAG) Chatbots for Education: A Survey of Applications," *Appl. Sci.*, vol. 15, no. 8, p. 4234, 2025, [Online]. Available: https://www.mdpi.com/2076-3417/15/8/4234

[8] N. M. Nalina Suresh, "Career Counseling Chatbot on Facebook Messenger using AI," *ACM Int. Conf. Proceeding Ser.*, 2021, [Online]. Available: https://dl.acm.org/doi/10.1145/3484824.3484875

[9] Monika D S, Dr. K. Balaji, "A Review on AI-Based Chatbots for Personalized Career Guidance," *Int. J. Adv. Res. Comput. Sci. Technol.*, vol. 8, no. 5, 2025, [Online]. Available: https://ijarcst.org/index.php/ijarcst/article/view/131

[10] Mashrur Rahman, Mantaqa abedin, "Transforming Mentorship: An AI Powered Chatbot Approach to University Guidance," *arXiv:2511.04172v1*, 2025.

[11] R. Lucien and S. Park, "Design and Development of an Advising Chatbot as a Student Support Intervention in a University System," *TechTrends 2023 681*, vol. 68, no. 1, pp. 79–90, Nov. 2023, doi: 10.1007/s11528-023-00898-y.

[12] Maria Eleni Koutsiaki, Marina Delianidi, Chaido Mizeli, Konstantinos Diamantaras, Iraklis Grigoropoulos, Nikolaos Koutlianos, "From Textbook to Talkbot: A Case Study of a Greek-Language RAG-Based Chatbot in Higher Education," *arXiv:2601.14265*, 2025, [Online]. Available: https://arxiv.org/abs/2601.14265