Final Project Proposal - CS557

Project Title: Advisor Recommendation AI System for BSU CS Graduate Students

Team Members:

1. Meherunnesa Tania - meherunnesatania@u.boisestate.edu - ID: 114275872
2. Miguel Carrasco Belmar - miguelcarrascobe@u.boisestate.edu - ID: 114286903
3. Afsana Afrin - afsanaafrin@u.boisestate.edu - ID: 114303243

Abstract:

Starting graduate school can be both exciting and overwhelming, especially for students coming from different countries and academic backgrounds. One of the biggest challenges new Computer Science graduate students at Boise State University face is finding the right research advisor. Many students struggle to understand faculty research areas, and simply do not know which professors are currently available to take new students. This often makes the advisor selection process stressful, uncertain, and time-consuming.

To address this challenge, our project introduces an AI-powered advisor recommendation system designed to make the process easier, more transparent, and equitable. The system uses Artificial Intelligence (AI) techniques such as Natural Language Processing (NLP) and Retrieval Augmented Generation (RAG) to understand students' questions in everyday language, retrieve key information about professors, like their research interests, publications, and availability, and generate helpful, context-aware responses.

Our system offers accurate and personalized guidance to help students find suitable advisors, reducing stress, building confidence, and ensuring equal access to information and opportunities. Beyond Boise State, this approach can guide other universities in improving support for their graduate communities through intelligent, student-focused tools. Overall, the project highlights how AI can make higher education more inclusive, approachable, and student-friendly.

Relevance to Course & AI:

This project directly connects to core AI concepts and the course's content. It will apply Natural Language Processing to understand user queries and generate coherent natural-language answers. It will also implement a Retrieval-Augmented Generation (RAG) framework, which we have studied this semester, to combine information retrieval with language generation. It will retrieve relevant information from a database of information about professors (contact information, research papers, area of study, etc.). By combining these techniques, the model will produce more accurate, relevant, and context-aware responses.

Project Significance:

The process of selecting a graduate advisor is a crucial step in a student's academic journey. At Boise State University, Computer Science graduate students are expected to secure a permanent advisor by the end of their second semester. As the official guidance states: *"The student should find a permanent advisor within the first semester, but no later than the second semester in the program. Students should talk to other faculty about their various research so as to find a topic that interests the student."* (Boise State University, 2025).[1]

However, this process can be difficult, especially for international students or those unfamiliar with the U.S. academic system. Many students arrive from diverse cultural and educational backgrounds, often facing language barriers, limited social networks, and unfamiliarity with faculty research areas. Research supports these challenges. Marijanović et al. (2021) [2] found that international doctoral students frequently feel stressed and uncertain while "speed-dating" with faculty members within limited timeframes to find a permanent advisor, especially since many of them are unsure of their research agendas and graduate school goals. These challenges can make it difficult to initiate conversations with professors, assess compatibility, and make decisions that will affect the future and focus of their education and professional development.

By developing this project, we aim to reduce stress and uncertainty for new graduate students by giving them personalized advice based on their skills and interests. We also seek to promote equity, allowing everyone to have access to the same information and guidance, no matter their background.

Beyond the Computer Department, this system could serve as a model for other departments and become a Boise State University standard. This aligns with the vision of Boise State GenAI's initiative of being the leader in Idaho for ethical, innovative, and accessible AI education that empowers students, faculty, and our community to shape the future responsibly [3].

Furthermore, this system could scale beyond BSU and benefit other institutions that are looking for ways to modernize their advisor systems in an effective and user-friendly way.

Novelty

Summary of Existing Work:

The concept of using recommender systems to assist students with academic choices is a well-established area in educational technology. Existing research primarily focuses on recommending courses, followed by university or degree program selection [4].

Previous models for academic recommendations have utilized various approaches:

- **Traditional Machine Learning and Filtering:** These include Collaborative Filtering (CF), Content-Based Filtering (CBF), and Hybrid Models combining the two or

integrating with methods like Decision Trees, Support Vector Machine (SVM), and K-Nearest Neighbor (KNN) [5].

- **Text-Based Similarity Models:** Systems have been proposed using methods like TF-IDF, Latent Semantic Analysis (LSA), and Word2Vec to calculate similarity for course recommendation systems, but these often lack semantic depth [6].

- **Deep Learning/Language Models (LLMs):** Recently, advanced models like BERT have been fine-tuned to extract text features and train sentence vectors for academic advisory recommendations, as demonstrated by the AdVisor Recommendation (AVRD) model. The AVRD model combines Chinese BERT with unsupervised Simple Contrastive Learning of Sentence Embeddings (SimCSE) and a time decay factor, showing superior performance over traditional models and general-purpose LLMs like Qwen and DeepSeek for the specific task of matching student interests to advisor profiles [7].

Identification of Gaps

Despite the progress in academic recommender systems, particularly models like AVRD, there are several key gaps that our project addresses:

- Existing recommendation systems mainly focus on course and program selection. Other existing work, like AVRD, primarily focuses on a general 'advisor' or 'teacher' recommendation. It is not a system tailored to a specific institutional context, like finding a research supervisor based on a graduate student's interests.

- Traditional academic recommendation systems provide a static, ranked list of matches based on keyword or vector similarity. They do not use a Retrieval-Augmented Generation (RAG) framework to handle natural language, conversational queries like "Who is available?" etc., and deliver a human-like, context-aware answer.

- Much of the important research, including the AVRD model, remains theoretical in nature, but often lacks deployment on a user-friendly, interactive platform for immediate use.

Project Contribution

The core contribution is providing a RAG-powered conversational AI solution for specific graduate advisor selection.

- We are introducing the Retrieval-Augmented Generation (RAG) framework into the graduate advisor recommendation domain. This is a crucial distinction from models that rely on pure similarity matching like AVRD, which uses BERT/SimCSE for vector generation or traditional LLMs without external context.

- By combining a search-efficient embedding model with an LLM (FLAN-T5 Base), our system will be able to understand nuanced, conversational queries like "Which professors are available for new students in Spring 2026?" etc, in a way that traditional vector similarity models cannot. Our system will retrieve heterogeneous, structured information

like availability, publications, etc, from a local database. It will also generate coherent, context-aware, natural language responses.

- Our recommendation system targets the critical, high-stakes task of securing a permanent research advisor, especially for students under time pressure. It addresses advisor-student compatibility by incorporating essential, real-world constraints like current faculty availability, which general recommendation models typically ignore.

Objective & Deliverables:

Objectives: The primary goal of this project is to design and implement an AI-powered advisor recommendation system that assists Boise State University Computer Science graduate students in finding suitable research advisors efficiently and equitably.

The specific objectives are to:

1. **Develop** an intelligent chatbot interface that allows students to ask natural-language questions about potential advisors ("Who is doing AI research?" or "Which professors are available for new students in Spring 2026?" or keywords, like Artificial Intelligence(AI), Machine Learning(ML), Computer Vision(CV)).
2. **Implement** a Retrieval-Augmented Generation (RAG) model that combines Natural Language Processing (NLP) with information retrieval to provide accurate, context-aware answers.
3. **Create and maintain** a structured professor database containing information such as faculty name, research areas, recent publications, contact details, and current availability for new students.
4. **Generate** personalized advisor recommendations based on students' interests, research keywords, and academic backgrounds.
5. **Design** a user-friendly web interface or chatbot prototype that BSU students can easily access and interact with.
6. **Evaluate** the system's performance and usability through user testing, focusing on accuracy, helpfulness, and ease of use.

Deliverables

By the end of the project, the team will deliver:

1. **Functional Prototype** – A working AI-based advisor recommendation system (chatbot or web interface) capable of responding to student queries using NLP and RAG.

2. **Faculty Information Database** – A well-structured dataset of BSU Computer Science faculty, including their research interests, keywords, and basic profiles.
3. **Recommendation Engine** – The backend component that processes user queries and generates personalized advisor matches.
4. **Evaluation Report** – A summary of testing results, including metrics such as response accuracy, relevance, and user satisfaction.
5. **Final Presentation & Documentation** – A comprehensive project report and demo presentation covering design choices, implementation details, ethical considerations, and future improvements.

Scope & Feasibility:

To keep the project focused and achievable, we have defined the scope of work:

In scope:

- Development of an AI-powered advisor recommendation system focused on Boise State University's Computer Science graduate programs.
- Integration of Natural Language Processing and Retrieval-Augmented Generation (RAG) to understand user input and provide context-aware responses.
- Creation of a database with faculty contact information, availability, projects, research interests, and areas.
- Implementation of a chatbot interface that allows students to interact with the AI-advisor system.

Out of Scope:

- Other BSU departments or graduate programs outside Computer Science.
- Non-academic advising topics (e.g., housing, financial aid, visa, or personal issues).
- Real-time integration with external university systems or official BSU domains.

Resources:

We are planning on using the following resources:

- BSU Computer Science faculty webpages for official professor information.

- Google Scholar for faculty research papers and publication data.
- Custom dataset (CSV) compiled manually for structured access.

Tools and Software:

- Python for backend development and AI model integration
- Hugging Face Transformers and SentenceTransformers libraries for NLP and embedding generation (open source)
- Streamlit for building an interactive web-based user interface (Free).
- Google Colab/VS Code for model development (Free).
- GitHub for version control and collaboration (Free).
- Borah Supercomputer for optional high-end model testing (already accessible with BSU accounts).
- Boisestate.AI and other AI assistants for troubleshooting and guidance.

Methods:

1. **Data Collection:** Data will be extracted from BSU websites, research repositories, and professors' individual websites. It will be cleaned and structured into a searchable format in a CSV file.
2. **Embedding Model Retrieval**: SentenceTransformers(all-MiniLM-L6-v2) will be used to encode both faculty profiles and student queries into vectors for similarity search.
3. **Language Generation Response:** FLAN-T5 Base to generate clear, context-aware text responses based on retrieved data.
4. I**ntegration and Interface:** Streamlit will be used to connect all components and allow students to interact with the system via chat-style queries.
5. **Evaluation**: We will evaluate the system using quantitative metrics (retrieval accuracy and response relevance) and qualitative feedback (direct user feedback among a few of our BSU graduate classmates)

Ethical Considerations:

The project prioritizes fairness, transparency, and privacy in developing the AI system.

1. **Data Privacy:** Only publicly available faculty information will be used. No personal or confidential data will be collected or stored.
2. **Fairness:** The system will ensure equal treatment of all users and avoid bias toward specific professors or research areas.
3. **Transparency:** Recommendations will include clear explanations of how matches are made. The tool will assist students but not replace personal communication with faculty.
4. **User Consent:** Students will be informed about the system's purpose and limitations before using it.
5. **Human Oversight:** Faculty or staff will review AI outputs to ensure accuracy and relevance.
6. **Data Accuracy:** Information sources will be regularly checked to prevent misinformation or outdated content.

Team Responsibilities:

1. Meherunnesa Tania
    - Responsibilities: Project Planning & Data Collection, Data Preprocessing & Embedding Setup.
    - Strengths: Strong analytical thinking skills and a solid understanding of data handling, and is proficient in Python, Pandas, and NumPy.

2. Miguel Carrasco Belmar
    - Responsibilities: RAG Pipeline Integration, Conversational Interface Development
    - Strengths: Strong programming background with experience in deep learning models, API integration, and system design.

3. Afsana Afrin
    - Responsibilities: Evaluation & Refinement, Technical Report & Finalization
    - Strengths: Excellent analytical and communication skills, with a strong background in interpreting machine learning results and writing technical reports.

Timeline:

Week	Milestone / Goal
Week 1: Project Planning & Data Collection	Define objectives and gather resources
Week 2: Data Preprocessing & Embedding Setup	Prepare data and retrieval foundation
Week 3: RAG Pipeline Integration	Combine retrieval and generation
Week 4: Conversational Interface Development	Build and connect the user interface
Week 5: Evaluation & Refinement	Test and improve system performance
Week 6: Technical Report & Finalization	Optimize and document the project

References:

[1] **Boise State University**, *Computer Science Masters' Student Handbook*, 2025. [Online]. Available: https://docs.google.com/document/d/11ktfPUE0AfkzvIQ22zAsvscSwvp49c8IPNz8lF21J9A/

[2] **L. Marijanović, N. Karraker, and M. Karpinska**, "*Advising experiences of first-year international doctoral students*," *Journal of International Students*, vol. 11, no. 4, pp. 940–959, 2021. [Online]. Available: https://files.eric.ed.gov/fulltext/EJ1300441.pdf

[3] **Boise State University**, *AI at Boise State – A culture of innovation*, 2025. [Online]. Available: https://www.boisestate.edu/genai/

[4] **N. Kamal, F. Sarker, A. Rahman, S. Hossain, and K. A. Mamun**, "*Recommender system in academic choices of higher education: A systematic review*," *IEEE Access*, vol. 12, pp. 35475–35501, 2024.

[5] **H. Ma, X. Wang, J. Hou, and Y. Lu**, "*Course recommendation based on semantic similarity analysis*," in *Proc. 3rd IEEE Int. Conf. Control Science and Systems Engineering (ICCSSE)*, 2017, pp. 638–641.

[6] **A. Baskota and Y.-K. Ng**, "*A graduate school recommendation system using the multi-class support vector machine and kNN approaches*," in *Proc. IEEE Int. Conf. Information Reuse and Integration (IRI)*, 2018, pp. 277–284.

[7] **X. Wang, J. Zhou, L. Jian, Y. Yin, and L. Li**, "*Empowering college students to select ideal advisors: A text-based recommendation model*," *Frontiers in Education*, vol. 10, 1673956, Oct. 2025.