



**BECS 32253 - Artificial Intelligence  
Project Proposal**

**AI Chatbot for Student Helpdesk**

STUDENT NO	NAME
EC/2021/001	K.P.T.T.D.S. Arshakularathna
EC/2021/060	B.R.M.D Baalasooriya
EC/2021/074	D.R.R.C Dasanayake
EC/2021/076	M.D.N Indeewari
EC/2021/080	G.U.G.K.U Jayarathne

## **1. Introduction and Background**

Students often face challenges when trying to access essential academic and administrative information, including examination schedules, course registration procedures, leave applications, and institutional guidelines. Traditional helpdesk services are frequently limited by long response times, restricted operating hours, and a high volume of inquiries, which can result in delays and frustration for students.

Advancements in Artificial Intelligence (AI), particularly Natural Language Processing (NLP), offer a solution to these challenges by enabling the development of intelligent chatbots. These systems can understand human queries expressed in natural language and provide accurate, context-aware, and instant responses.

The availability of real-world datasets, such as DSSC-Final-Datasets.jsonl, containing question answer pairs related to student services and academic processes, allows the creation of a data-driven AI chatbot. By training the chatbot on this dataset, the system can automatically respond to common student inquiries, reducing the workload of administrative staff, providing 24/7 support, and ensuring consistent, reliable assistance.

This approach results in a smart virtual helpdesk that enhances communication between students and the institution, improves operational efficiency, and contributes to a more accessible and responsive academic environment.

## **2. Problem Statement**

Students frequently face challenges in accessing clear and accurate academic and administrative information, such as examination schedules, registration guidelines, leave procedures, and institutional rules. The required information is often distributed across multiple platforms, including websites, notice boards, emails, and verbal instructions, making it difficult and time-consuming to locate. As a result, students may experience confusion, misunderstandings, or delays in completing important academic tasks.

At the same time, administrative staff are consistently required to answer the same routine inquiries, leading to repetitive work and increased workload, especially during busy periods such as registration weeks or examination sessions. This creates bottlenecks in communication, reduces service efficiency, and can negatively affect the student's experience.

Therefore, there is a need for a centralized, reliable, and continuously available system that can provide quick and accurate responses to common student questions. Developing an AI-powered chatbot using Natural Language Processing (NLP) and real institutional Q&A data offers an effective solution. Such a system can operate 24/7, reduce staff workload, ensure consistent delivery of information, and enhance communication between students and the institution.

### **3. Main Objectives**

The primary objectives of this project are to design, develop, and evaluate an intelligent chatbot capable of assisting students with their inquiries efficiently. Specifically, the project aims to:

1. Develop an AI-powered chatbot that can automatically understand and respond to student-related queries in natural language.
2. Integrate Natural Language Processing (NLP) techniques to accurately interpret and classify student questions, enabling meaningful and context-aware responses.
3. Train the chatbot using a real dataset of question answer pairs derived from the institution's frequently asked questions (FAQs) or relevant academic resources.
4. Design and implement a user-friendly web-based interface that allows students to interact seamlessly with the chatbot.
5. Evaluate the performance of the chatbot in terms of response accuracy, relevance, and user satisfaction, using real student queries as test cases.

These objectives ensure that chatbot is not only functional but also effective in improving communication, reducing administrative workload, and providing students with timely and reliable support.

### **4. Proposed Methodology**

<b>Step</b>	<b>Description</b>
<b>Data Collection</b>	Use the provided dataset (DSSC-Final-Datasets.jsonl) containing student inquiries and official responses.
<b>Data Preprocessing</b>	Clean text, remove unwanted characters, normalize text, and separate into Question - Answer pairs.
<b>Model Selection</b>	Use a <b>NLP model</b> such as DistilBERT / BERT fine-tuned for Question-Answer matching or a Retrieval-Based Chatbot using sentence embeddings (e.g., SBERT).
<b>Training</b>	Train or fine-tune the model to match student questions with the correct answers.
<b>Chatbot Framework</b>	Implement using Python + Flask/Streamlit or integrate into Web UI.

<b>Response Generation</b>	When a user sends a query, the chatbot searches for the most similar question in the dataset and returns the corresponding answer.
<b>Evaluation</b>	Evaluate performance using metrics such as accuracy, response relevance, and user feedback.

## System Architecture Overview:

User → Chat Interface → NLP Model (Embedding + Similarity Matching) → DSSC Q&A Database → Response

## 5. Expected Outcomes

The project is expected to deliver a fully functional AI-powered chatbot capable of accurately responding to student queries in real time. The system will include a clean, user-friendly web application that allows students to interact with the chatbot easily. The chatbot will be supported by a trained NLP model developed using the provided institutional dataset, ensuring relevant and context-aware responses. Additionally, a performance evaluation report will be prepared based on accuracy testing and user feedback to assess the system's effectiveness. Finally, full project documentation and a demonstration will be provided to showcase the system's features, workflow, and practical usability.

## 6. References

- Dataset: *DSSC-Final-Datasets.jsonl*
- Jurafsky, D., & Martin, J. H. (2022). *Speech and Language Processing*.
- Devlin, J., et al. (2018). *BERT: Pre-training Deep Bidirectional Transformers for Language Understanding*.
- Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence Embeddings for Semantic Similarity*.