

CMP6200

Individual Undergraduate Project

2024 - 2025

A1 - Proposal

University Artificially Intelligent Assistant



Course: Computer & Data Science
Student Name: Lewis Higgins
Student Number: 22133848
Supervisor Name: Dr. Atif Azad

Contents

1	Introduction	2
1.1	Background and Rationale	2
1.2	Key Themes/Topics	2
2	Aims and Objectives	3
2.1	Project Aim	3
2.2	Project Objectives	3
3	Project Planning	4
3.1	Initial Project Plan	4
3.2	Resources	4
3.3	Risk Assessments	5
4	Project Review and Methodology	6
4.1	Critique of Past Similar Projects	6
4.2	Literature Search Methodology	6
4.3	Initial Literature Search Results	6
	Bibliography	6

Introduction

1.1 Background and Rationale

With artificial intelligence (AI) becoming increasingly more powerful and useful in recent years, in some cases even surpassing humans in areas like language and image recognition (Giattino et al., 2023). It is therefore essential that higher education institutions take advantage of it and ensure to keep up-to-date with its developments in the interest of academic integrity, which was also stated in the Higher Education Policy Report dated 28 March 2024 - "Higher education will have to adopt, adapt, collaborate and lead to take advantage of AI while managing the risks" (HEPI, 2024).

This project aims to leverage recent technical developments in natural language processing (NLP) to create a digital assistant for university life that students can use to gain information on topics such as university policies and locations on campus. This is because attending university for the first time is a daunting and stressful experience for many, often due to it being a new and unfamiliar environment where students have full independence unlike their previous educational settings, which could possibly lead to declines in academic performance and social activity. Some students may not wish to speak with newer people who they don't know about university topics for fear of ridicule or embarrassment, and would benefit from a digital companion to help them to become acquainted with their new environment without the perceived risk of social judgement.

Chatbots are also a significant tool across the tech sphere, with 73% of surveyed web users expecting companies to have chatbots for convenient interactions (Cherniak, 2024). By giving students a quick and easy tool to get the information they need, they can instead shift their focus to their studies instead of being distracted by smaller issues.

1.2 Key Themes/Topics

This project undertakes the following key themes:

- Natural Language Processing (NLP)
 - As the backbone of this project, extensive research into this topic will be necessary to ensure users have a smooth experience.
- Embedding models
 - To store non-numeric data such as university policies, a suitable embedding model will be necessary to vectorise said data into a numerical representation interpretable by the machine learning model.

Aims and Objectives

2.1 Project Aim

This project aims to aid new and existing students alike while they are attending university with helpful information about the university itself, such as university societies, locations/campuses, and policies through the medium of a digital chatbot companion.

2.2 Project Objectives

- WIP.

Project Planning

3.1 Initial Project Plan

1. Research
 - (a) Conduct heavy research into machine learning and natural language processing to bolster my knowledge of the topics to assist in the development of the chatbot.
 - (b) Identify similar projects that already exist to understand where challenges may arise in development and how to differentiate my work to make it stand out and provide unique value.
2. Data collection
 - (a) To present users with information from the university, I must first collect this information for myself from sources such as the university website and the Student Union.

3. Section WIP.

3.2 Resources

- An integrated development environment (IDE) for Python
 - Visual Studio Code is a lightweight editor that supports most programming languages, including Python.
 - An alternative could be JetBrains' PyCharm Professional, which I can access at no charge due to being a student.
- Machine learning libraries & frameworks
 - Examples include PyTorch, TensorFlow and SpaCy.
- A powerful computer.
 - Training machine learning models requires significant processing power and RAM. I own a decently powerful computer with a higher-end graphics card which should be able to handle a project of this scale.
- A platform for the chatbot.
 - Many messaging services allow developers to add bots, such as Facebook Messenger, WhatsApp or Discord.

3.3 Risk Assessments

Risk	Likelihood	Severity	Overall	Mitigation
The devices used to write code and documentation could encounter hard drive corruption, potentially losing considerable amounts of work.	2	5	Medium	Ensure that all work is regularly backed up to the cloud and/or a secondary device. Github will be used as a version control system for the project to keep an audit trail of all changes.
Time constraints could potentially cause rushed and poor-quality work if development is not to a consistent and regular schedule balanced with other university modules.	3	4	Medium-High	Ensure that work is completed at regular intervals, even if the amount at each interval is small. In doing so, it will be much easier to balance three simultaneous deadlines.
Personal circumstances could cause progress to slow or halt if something unexpected were to occur during the year.	3	4	Medium-High	Try to be ahead of deadlines where possible to ensure that there is free time that could be used in the event of a sudden break becoming necessary.
Budget constraints could be an issue during development.	2	2	Low	Use open-source or lower cost resources during development, and create a budget to adhere to.

Project Review and Methodology

4.1 Critique of Past Similar Projects

While researching similar projects, I identified three useful final year projects by BCU students. The first was a 2019 project by Sanah Mehreen Hussain, which was the creation of a chatbot to assist students with their module content. As another chatbot project, the processes they undertook in their development will be heavily interlinked with my own, which made it imperative to review. Their report was highly detailed, providing a clear and strong explanation of each step of their development process. They had created many diagrams to clearly explain their processes, making it a powerful learning resource. Their chatbot was placed within Moodle itself, which is used by BCU students to access all of their work, which would massively improve its engagement. Sanah also conducted surveys for user feedback which proved useful in their evaluation.

Another useful project to review was Ali Akbar Rashid's 2022 chatbot for BCU IT support. This was another highly detailed exploration of the chatbot development process. Ali used the Agile project management methodology in the development of their project, citing its fast approach and lower cost in relation to other methodologies such as Waterfall, as well as frequent opportunities for feedback. Even still, Ali mentioned time limitations due to the balancing of his project with his other university modules, which suggests that I must take particular care to ensure I am setting and meeting frequent goals as mentioned in Section 3.3. Ali's evaluation and research was also heavily detailed, with a strictly defined methodology for both.

The third useful project was "KURA", a chatbot to assist users with information about the 2022 World Cup, developed in 2023 by Stanley Eweniwe Osuozah. While it is not as directly linked to a university chatbot like the others, it is functionally similar in that it is also retrieving information and displaying it to users, though simply of a different topic. This project was not documented quite as well as the other two, though it did still have some useful information to be learned. Their user survey was somewhat confusing to interpret, with statements with different implications such as "the KURA chatbot is welcoming" followed immediately by "the KURA chatbot is not clear about what it is meant for." This shift from a positive statement to a negative one within the survey could confuse participants and give potentially misleading results due to their misinterpretation of what they are being asked. The dataset they published to Kaggle was also very limited, and did not yield much useful information. From this project in particular, I have gathered that I must convey my intentions clearly and in significant detail to ensure that my work could be used by others in future to assist the development of their own projects.

4.2 Literature Search Methodology

AAAAAA

4.3 Initial Literature Search Results

AAAAAA

Bibliography

- Giattino, C., E. Mathieu, V. Samborska and M. Roser (2023). *Artificial Intelligence*. URL: <https://ourworldindata.org/artificial-intelligence?insight=ai-systems-perform-better-than-humans-in-language-and-image-recognition-in-some-tests#key-insights> (visited on 08/10/2024).
- HEPI (28th Mar. 2024). *New HEPI Report explores the impact of technology on universities and students*. URL: <https://www.hepi.ac.uk/2024/03/28/technology-foundations-for-twenty-first-century-higher-education/> (visited on 08/10/2024).
- Cherniak, K. (22nd Sept. 2024). *Chatbot Statistics: What Businesses Need to Know About Digital Assistants*. URL: <https://masterofcode.com/blog/chatbot-statistics> (visited on 09/10/2024).