A Chatbot for a Study Program at UiT

Engineering Student Report
2nd Year Internship
Academic Year 2023-2024

Presented by:
Roubille Mathis

Academic Advisor:
Viet Hung NGUYEN

Internship Supervisor:
Hao YU

Duration of the internship: 17 weeks

August 23, 2024

Acknowledgments

I would like to express my deep gratitude to all the people who supported and guided me throughout my internship at the Arctic University of Norway (UiT).

First of all, I sincerely thank my internship supervisor, Hao YU, for his exceptional guidance, availability, and valuable advice. His expertise and support were essential to the success of this project.

I also wish to thank Viet Hung NGUYEN, my academic advisor at ISIMA, for his follow-up. His academic guidance was of great help to me in completing this internship.

Finally, I would like to express my gratitude to all the members of the Industrial Engineering department at UiT for their warm welcome and collaboration. Their help and encouragement greatly facilitated my integration and enriched my experience.

I thank you all for your trust and support throughout this academic and professional adventure.

Contents

Acknowledgments I

Résumé V

Abstract V

Glossary VI

1 Introduction 1

2 UiT the Arctic University Norway 3

3 Planning 5

4 Stakeholder Needs Analysis 6
4.0.1 Test and Interview Preparation 6
4.1 Analysis Results 6
4.2 Implications for Chatbot Development 7

5 Technology Choice for Chatbot Development 8
5.1 Introduction 8
5.2 Intent Detection Models 8
5.3 Use of Large Language Models (LLM) 9
5.4 Conclusion 9

6 Functioning 10
6.1 Introduction 10
6.2 Question Processing 10
6.2.1 Contextual Enrichment 11
6.3 Response Formatting and Conversion 11
6.3.1 Markdown to HTML Conversion 11
6.4 User Interaction Examples 12
6.5 Technical Challenges and Solutions 12

7 Testing 14
7.1 Test Methodology 14
7.2 Test Results 15

7.3 Analysis of Results and User Feedback 16

8 Integration into the Website **17**

8.1 Popup Interface 17

8.2 Key Features 17

8.3 Attention Animation 17

8.4 Integration Objectives 17

9 Simulation **19**

9.1 Simulation Environment Description 19

9.2 Simulation System Parameters 19

9.3 Simulation Scenarios and Results 20

9.4 Analysis of Simulation Results 21

9.4.1 Comparison with a System without a Chatbot 21

9.4.2 Impact of the Chatbot (50.00% Intelligence) 21

9.4.3 Increasing Chatbot Intelligence 22

9.4.4 Reduction of Human Workload 22

9.4.5 Reduction of Required Human Agents 22

10 Conclusion **23**

Références bibliographiques (norme IEEE) **26**

List of Figures

1 UiT Campus . . . 3
2 Hierarchical Organization at the University of the Arctic of Norway (UiT) . . . 4
3 Gantt Chart . . . 5
4 Chatbot Flowchart . . . 10
5 Chatbot Test Interface . . . 14
6 Distribution of User Ratings . . . 15
7 UiT Chatbot Interface Closed . . . 18
8 UiT Chatbot Interface Open . . . 18

Résumé

Ce projet de stage a porté sur le développement et l'intégration d'un chatbot pour le programme de Master en Ingénierie Industrielle de l'Université de Narvik (UiT) en Norvège. L'objectif était de créer un outil capable de fournir des réponses automatiques et contextuellement appropriées aux questions des étudiants potentiels, améliorant ainsi leur accès à l'information et réduisant la charge de travail des agents humains. Pour ce faire, j'ai utilisé l'API GPT-3.5 Turbo, un grand modèle de langage (LLM), pour générer des réponses précises. Le processus de développement a inclus une analyse approfondie des besoins des parties prenantes, la conception et la mise en œuvre du chatbot, ainsi que des tests rigoureux pour évaluer son efficacité. Les résultats des tests ont montré une satisfaction élevée des utilisateurs, avec une réduction notable des temps de réponse et des abandons. La simulation réalisée avec AnyLogic a confirmé l'efficacité du chatbot, démontrant une amélioration significative par rapport aux solutions précédentes. En intégrant cette technologie, l'UiT peut offrir une expérience utilisateur améliorée et une meilleure allocation de ses ressources humaines.

Mots-clés : Chatbot, Ingénierie Industrielle, UiT, Norvège, GPT-3.5 Turbo, LLM, Simulation, AnyLogic.

Abstract

This internship project focused on the development and integration of a chatbot for the Master's program in Industrial Engineering at Narvik University (UiT) in Norway. The goal was to create a tool capable of automatically providing contextually appropriate answers to potential students' questions, thereby improving their access to information and reducing the workload of human agents. I used the GPT-3.5 Turbo API, a large language model (LLM), to generate precise responses. The development process included a thorough analysis of stakeholder needs, the design and implementation of the chatbot, and rigorous testing to evaluate its effectiveness. Test results showed high user satisfaction, with notable reductions in response time and drop-offs. Simulation conducted with AnyLogic confirmed the chatbot's effectiveness, demonstrating significant improvements over previous solutions. By integrating this technology, UiT can offer an enhanced user experience and better allocation of its human resources.

Keywords: Chatbot, Industrial Engineering, UiT, Norway, GPT-3.5 Turbo, LLM, Simulation, AnyLogic.

Glossary

AnyLogic Discrete event modeling software used to simulate the arrival flow of users in a customer service system. 19, 20, 23, V

API Application Programming Interface, a set of functions and procedures allowing the creation of applications accessing the functionalities or data of another service or software. 9–11, 13, 23, V

BERT Bidirectional Encoder Representations from Transformers, a pre-trained language model based on transformers, effective in capturing contextual relationships between words in a sentence. 8

CAO/FAO Computer-Aided Design / Computer-Aided Manufacturing, techniques used to design and manufacture products with the aid of computer software. 3

ELIZA A natural language processing computer program developed in the 1960s by Joseph Weizenbaum at MIT, designed to simulate a natural language conversation. ELIZA was one of the first chatbot programs and demonstrated the possibilities and limitations of human-machine interactions. 1

GPT Generative Pre-trained Transformer, a language model architecture developed by OpenAI, capable of generating coherent and contextually relevant text based on a large number of input parameters. Advanced versions of GPT, such as GPT-3.5 and GPT-4, are widely used in applications ranging from virtual assistance to creative content generation. 1

GPT-3.5 Turbo Language model developed by OpenAI, used to generate contextually appropriate responses in the development of the chatbot. 7, 9–13, 23, V

HTML HyperText Markup Language, the standard language used to create and structure web pages. 11, 13

LLM Large Language Model, a large language model capable of generating coherent and relevant text based on the given context. 1, 9, 23, II, V

Markdown Lightweight markup language allowing simple and readable text formatting. 11, 13

NLP Natural Language Processing, technology used for developing chatbots and other text processing applications. 1, 2, 23, 24

RNN Recurrent Neural Network, a recurrent neural network capable of capturing temporal dependencies in data. 8

SVM Support Vector Machines, a type of classifier used for intent detection in chatbot models. 8

UiT University of the Arctic of Norway (Arctic University of Norway), located mainly in northern Norway. 1, 3, 6, 8, 9, 11, 14, 16, 17, 23, 24, I, V

1 Introduction

The rapid evolution of natural language processing (NLP) technologies and chatbots has profoundly transformed the way businesses and institutions interact with their clients and users. From the early rudimentary language processing systems in the 1960s, such as ELIZA[1], to the sophisticated models of today like GPT[2], the progress has been substantial. These innovations have enabled the creation of systems capable of understanding and generating natural language with impressive accuracy and fluency. This transformation extends to various fields, including education, where integrating these advancements can offer new support and interaction tools. Chatbots can respond in real-time to a multitude of questions, provide personalized information, and guide users through complex processes, thereby enhancing the overall experience for students and stakeholders. My internship project aims to develop a chatbot specifically tailored to the needs of students interested in the Master of Industrial Engineering program at the University of Narvik - the Arctic University of Norway (UiT).

The main problem this project seeks to address is the lack of immediate and accessible support for prospective students seeking information about the Master of Industrial Engineering program. Currently, students and stakeholders often have to navigate through disparate sources of information or wait for responses to their questions, which can discourage some candidates. Therefore, the objective is to create a chatbot capable of automatically answering a predefined range of questions related to this study program, thereby improving the user experience and facilitating access to information. Additionally, this project aims to reduce the human workload required to answer these questions, allowing for a better allocation of human resources.

This project continues from a previous attempt at chatbot simulation, which did not yield fruitful results. About 50% of users of this experimental system reported dissatisfaction with the chatbot services[3]. This previous experience underscores the importance of thorough analysis and careful design to effectively meet user expectations. However, the real difference between the first chatbot integration study at UiT and today lies in the rapid evolution of technologies and the explosion of large language models (LLMs). Modern LLMs, such as ChatGPT, offer much more advanced natural language processing capabilities and better contextual understanding, enabling the creation of much more efficient and effective chatbots.

The chatbot development process will involve several key stages, structured to ensure a methodical and effective approach. First, a thorough stakeholder analysis will be conducted to identify the needs and expectations of potential users. This phase is crucial to ensure the chatbot meets the actual requirements of the users. This analysis will include interviews, surveys, and the examination of historical data to understand frequently asked questions and current challenges faced by users. Next, we will proceed with the design and implementation of a chatbot prototype, incorporating stakeholder feedback to maximize its relevance and utility. The prototype will be developed using advanced NLP technologies and tested in realistic scenarios to evaluate its performance and usability. Once the prototype is developed, rigorous testing will be conducted with a representative sample of potential users to assess the effectiveness and usability of the solution. The data and feedback collected during this phase will be used to iterate and improve the chatbot. Finally, a simulation will be conducted to evaluate the chatbot's performance under near-real conditions, using simulation tools to model user interaction and predict large-scale outcomes.

This project offers a unique opportunity to acquire and develop a range of multidisciplinary skills, from stakeholder analysis to project management, chatbot development, and performance evaluation using simulations. By developing this chatbot, we aim to create a powerful and effective tool that not only meets the current needs of students and stakeholders but can also evolve to adapt to future challenges and opportunities in the field of education.

2 UiT the Arctic University Norway

The University of the Arctic of Norway (UiT) is among the leading academic institutions in the country, offering a wide range of high-quality educational and research programs. Located mainly in northern Norway, UiT has eleven campuses spread across cities such as Narvik where I was located but also Tromsø, Alta, and Harstad, creating a diverse and stimulating learning environment. [4]

Figure 1: UiT Campus

The Master of Science in Industrial Engineering program at UiT is designed to equip students with the knowledge and skills necessary to use appropriate techniques, tools, and methodologies to identify, formulate, analyze, and solve complex engineering and management problems. This program covers various fields, including CAO/FAO, industrial robotics, optimization and data management, virtual and integrated manufacturing, and supply chain management.

During my internship, I had the honor of working at UiT under the supervision of Hao Yu, a member of the Department of Industrial Engineering. Hao Yu is affiliated with the Intelligent Manufacturing and Logistics research group, which aims to become a national and international leader in scientific fields such as robotics, logistics, automation, production engineering, additive manufacturing, human-machine interaction, and human-robot collaboration.

This research group is at the forefront of strategic technology development. In a region characterized by large distances and a harsh climate, innovative technological solutions are essential for providing welfare services to residents. The group contributes to advancing knowledge on technological solutions that promote inclusive social development and diversification of economic activities in the north. It also focuses on technologies addressing health, external environment, and safety challenges, as well as digital competence in education.

During my internship, I worked autonomously on the project of developing a chatbot to answer questions about the Master's in Industrial Engineering, benefiting from Hao Yu's expertise and guidance throughout the process.

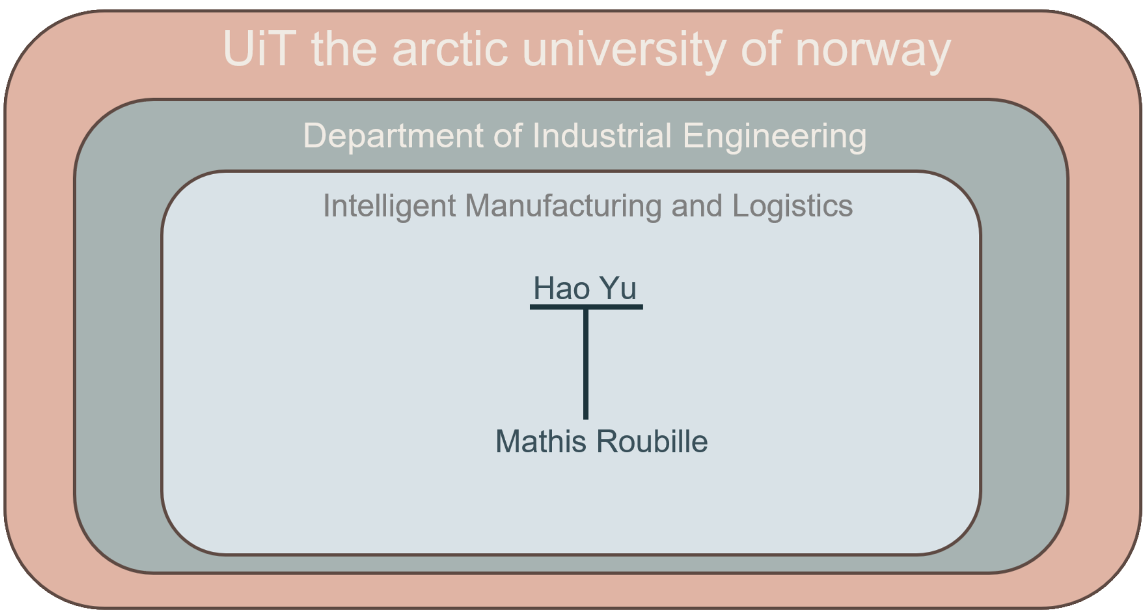

Figure 2: Hierarchical Organization at the University of the Arctic of Norway (UiT)

3 Planning

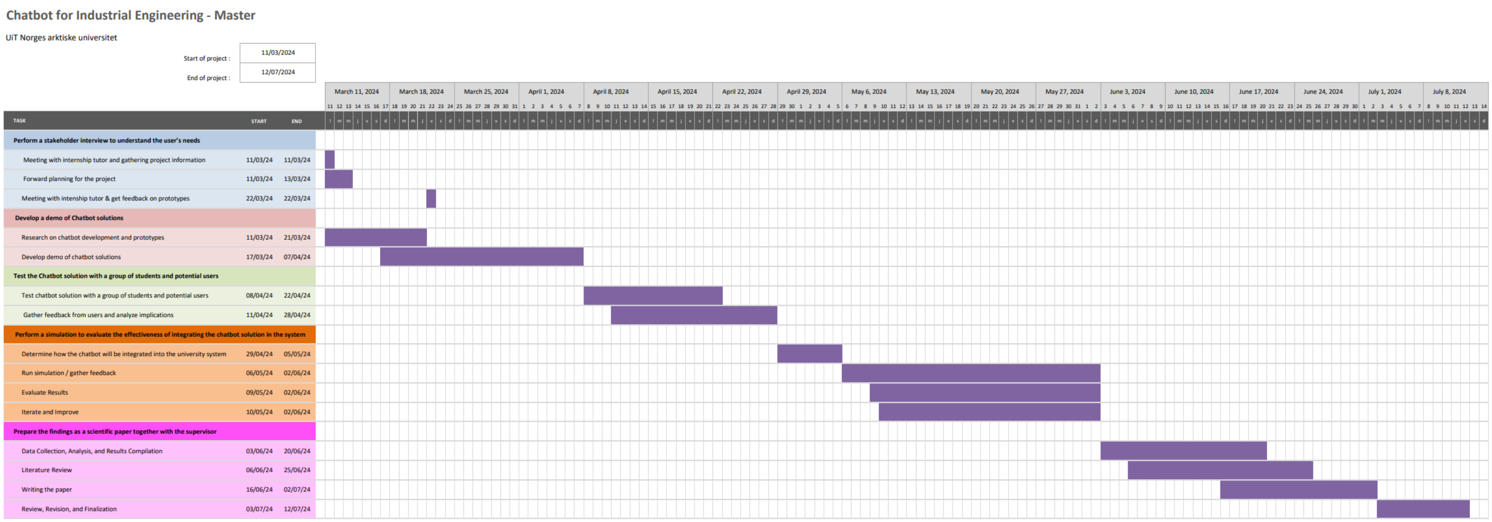

Figure 3: Gantt Chart

The project followed a planning divided into five main phases. The first phase was dedicated to conducting stakeholder interviews to understand their needs. This step was followed by developing a demonstration of chatbot solutions. Next, we tested the chatbot solution with a group of students and potential users to gather feedback and analyze the implications. Subsequently, a simulation was conducted to evaluate the effectiveness of integrating the chatbot into the university system. Finally, the last stage involved compiling and presenting the results in the form of a scientific paper, in collaboration with my supervisor.

The project progressed almost exactly as planned according to the Gantt chart. The Gantt chart was carefully crafted to detail the various stages of the project and the associated deadlines. Thanks to rigorous planning and disciplined execution, each phase was completed within the allocated time frame, allowing for constant alignment with the initially established schedule.

4 Stakeholder Needs Analysis

Analyzing stakeholder needs is a crucial step in developing any system. For this project, in-depth interviews were conducted with various actors, including students and administrative staff at UiT. These interviews provided valuable insights into the expectations and specific needs of the users. This section details the analysis process, the identified needs, and the implications for chatbot development.

4.0.1 Test and Interview Preparation

During the interviews, stakeholders were invited to interact with the current chatbot. They were encouraged to ask typical questions they would receive or ask about the Master's program in Industrial Engineering at UiT. After this interaction, structured interviews were conducted to discuss their experience.

- **Students**: Students tested the chatbot by asking questions about courses, admissions, student services, housing, social aid...
- **Administrative staff**: Administrative staff tested the chatbot on questions related to administration, student management, research groups, course content, available resources...

Expressed Needs	Priority
Quick and accurate responses to academic questions	High
Reduction of workload related to recurring questions	Medium
Improvement of the efficiency of information services	High

Table 1: Summary of Stakeholder Needs

4.1 Analysis Results

The interview results revealed several key needs among different stakeholders:

- **Students**: Students expressed the need to receive quick and accurate responses to their academic questions. They also want to access this information outside office hours for greater flexibility.
- **Administrative staff**: The administrative staff highlighted the need to reduce the workload related to managing recurring student questions. They prefer the chatbot to handle frequently asked questions, allowing them to focus on more complex and pedagogical tasks

and a need to improve the efficiency of information services for potential and current students. A chatbot capable of handling common requests can free up time for more critical administrative tasks.

4.2 Implications for Chatbot Development

The information gathered during the interviews was essential to guide the chatbot's development. They helped identify priority features and ensure the chatbot would meet the end-users' expectations. Here are some key implications for development:

- **Knowledge base**: An exhaustive knowledge base was developed to cover frequently asked academic questions by students. This knowledge base is regularly updated to stay relevant.
- **Contextual responses**: The chatbot was designed to provide precise contextual responses using advanced language models such as GPT-3.5 Turbo.
- **Workload reduction**: By automating responses to recurring questions, the chatbot helps reduce the workload of professors and administrative staff, allowing them to focus on more strategic tasks.

The methodology employed for the chatbot's design included several key steps: data collection, requirements analysis, development, and testing. All these steps were carried out in close collaboration with stakeholders to ensure the identified needs were well understood and addressed.

In conclusion, analyzing stakeholder needs played a crucial role in the project's success. By understanding the specific expectations of end-users, the chatbot's development was directed to effectively meet their needs, ensuring maximum adoption and satisfaction.

5 Technology Choice for Chatbot Development

5.1 Introduction

This section of the report examines the various technologies used in developing AI-powered chatbots. The primary objective is to select an appropriate technology based on the specific needs of the project, considering the advantages and challenges associated with each approach.

5.2 Intent Detection Models

Initially, I considered using an intent detection model for the chatbot. This type of model requires training on annotated data, where each question is associated with a predefined intent and several possible answers [5].

Different intent detection models can be used, ranging from simple classifiers like support vector machines (SVM) to more complex models like recurrent neural networks (RNN) or transformers [6]. SVMs are generally simple to implement but may lack precision in more complex cases. RNNs, on the other hand, are capable of capturing temporal dependencies in data, which can be beneficial for intent detection in longer or more complex dialogues. Transformers, particularly pre-trained language models like BERT, can also be effective as they capture contextual relationships between words in a sentence, which is crucial for understanding the meaning of questions [7].

However, intent detection models have their limitations. They require a large amount of annotated data covering various ways of asking a question and information about specific questions asked. Even with adequate training, it is difficult to ensure all responses will be satisfactory to users, as these are predefined in advance. Finding an exact match is rare, and it is nearly impossible to effectively cover multiple languages, which is crucial for an institution like UiT that welcomes international students.

5.3 Use of Large Language Models (LLM)

This is why I opted for the use of a large language model (LLM), specifically the API GPT-3.5 Turbo. An LLM works by generating coherent and relevant text based on the given context. The model is pre-trained on a vast amount of textual data from the internet, allowing it to have a deep understanding of natural language and generate contextually appropriate responses [2]. An LLM can be seen as an improved version of the word prediction systems found on our phones.

There are other LLM options outside of GPT-3.5 Turbo. For example, Mistral is a French LLM that, in its smaller versions, can operate locally while providing response quality comparable to GPT-3.5 Turbo. However, implementing Mistral would have required more work and research, without necessarily significant cost savings. LLMs in general, including GPT-3.5 Turbo and Mistral, represent significant advances in natural language processing, but their effectiveness often depends on various factors such as model size, training data quality, and available processing capacity.

A notable advantage of GPT-3.5 Turbo is its affordable economic cost, with competitive rates for predictions, making it particularly attractive for large-scale applications like the chatbot envisaged for UiT.

Despite the existence of GPT-4 at the time of the initial choice, I decided not to opt for this version. GPT-3.5 Turbo already offered satisfactory performance for the chatbot's needs, and the available access keys for GPT-3.5 Turbo facilitated initial development and testing.

A potential issue with using LLM like GPT-3.5 Turbo is the possibility for the chatbot to generate incorrect or unverified responses, as the model primarily aims to provide coherent information rather than accurate. However, after many tests, the chatbot never gave a response not based on the provided information, and it correctly indicated when it lacked the necessary information. Moreover, the chatbot is designed to answer only questions within the provided context to prevent its use for off-topic questions [5].

5.4 Conclusion

In conclusion, choosing an LLM like GPT-3.5 Turbo for chatbot development is based on its ability to provide precise and contextually appropriate responses in various languages while offering an affordable economic cost for predictions. Despite some potential limitations, this model represents a significant advancement in natural language processing for chatbot applications.

6 Functioning

6.1 Introduction

This section explores the detailed functioning of the developed chatbot, focusing on the process of handling user questions and generating responses via the API GPT-3.5 Turbo.

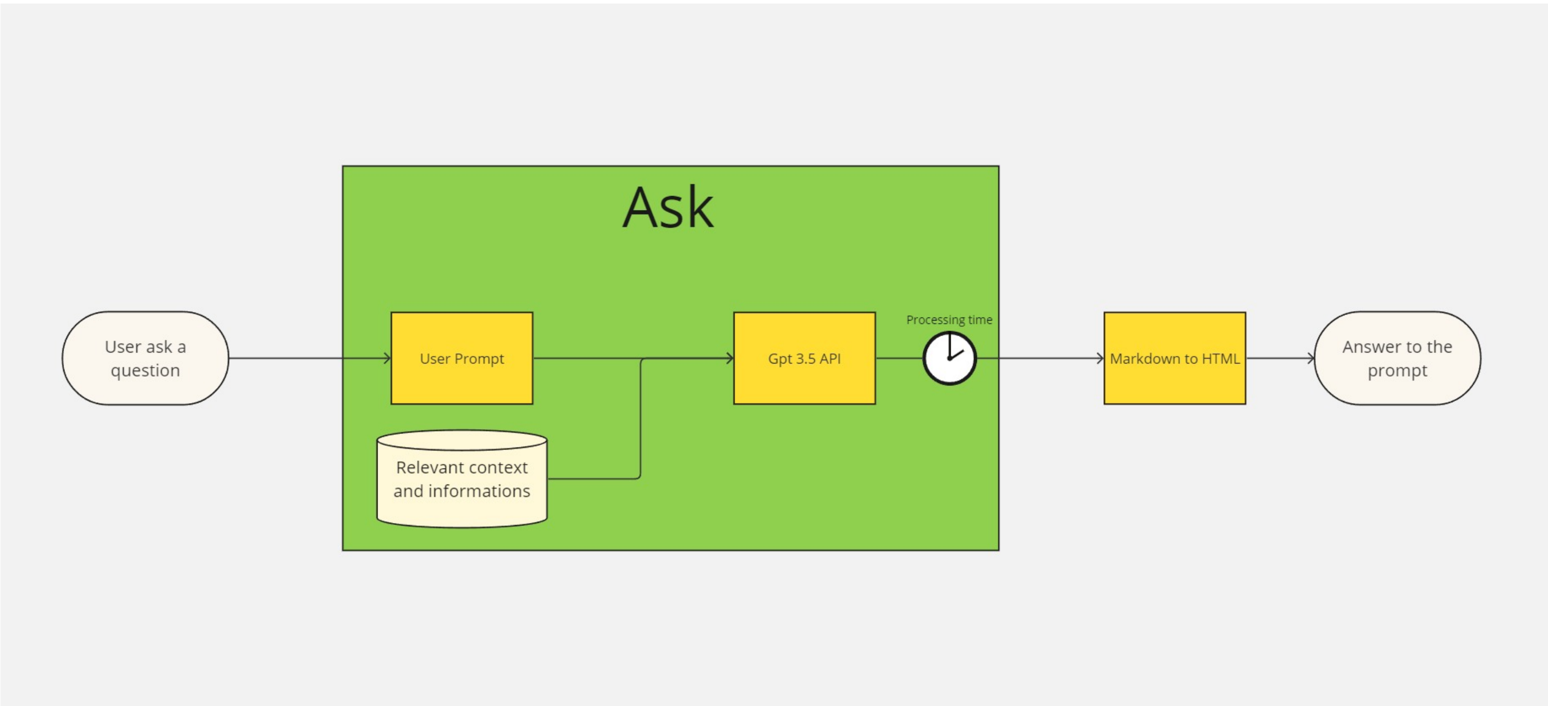

Figure 4: Chatbot Flowchart

6.2 Question Processing

The primary role of the chatbot is to act as an intermediary between the user and the API GPT-3.5 Turbo, facilitating communication and processing questions and answers. When receiving a user's question, the chatbot performs several processing steps. First, it extracts the raw text of the user's question. Then, it enriches this question by providing it with the context of the current conversation, as well as additional information about the degree, Narvik University, and other relevant topics. This data is essential for the API GPT-3.5 Turbo to have an updated and accurate knowledge base on the subject addressed.

6.2.1 Contextual Enrichment

After extraction, the chatbot enriches the question with the context of the current conversation and additional information. This includes details such as previous questions asked in the same session, as well as specific information about the programs, courses at UiT, student life, housing, access to Narvik, and other relevant topics.

The chatbot's information base mainly comes from the UiT website, where data about the Industrial Engineering program in Narvik and other related topics are available. This database was then supplemented with information provided by students and other people at UiT, as well as my own experiences at UiT and in Narvik. I ensured that this knowledge base was as complete as possible by adding relevant details and updating the information.

Once the user's question is enriched with the context and additional information, the chatbot sends it to the API GPT-3.5 Turbo to obtain a response. The response generated by the API GPT-3.5 Turbo is then returned to the chatbot.

6.3 Response Formatting and Conversion

The response provided by the API GPT-3.5 Turbo is in Markdown format. Markdown is a lightweight markup language that allows for simple and readable text formatting by adding tags like bold, italic, or creating bullet lists. However, to be displayed correctly on a web page or in an application, the response must be converted to HTML format.

6.3.1 Markdown to HTML Conversion

The process of converting Markdown to HTML is crucial to ensure that responses are displayed consistently and aesthetically on the UiT website. This process includes converting various Markdown elements to corresponding HTML tags while avoiding potential code injection risks.

To prevent any possibility of code injection, several precautions are taken. First, the Markdown markup characters are carefully replaced with their equivalent HTML tags securely. For example, double asterisks or underscores used for bold text are converted to ‘‘ tags, while single asterisks or underscores for italics are converted to ‘<em>‘ tags.

Additionally, hyperlinks in Markdown, represented by ‘[link text](URL)‘, are converted to ‘<a>‘ tags in HTML. Each link is verified to ensure it does not contain malicious content. HTML tags are generated securely to prevent any execution of injected code.

Here is an example of the conversion:

- **Markdown**: ‘**This course is required**‘
- **HTML**: ‘This course is required‘

This conversion process ensures that the chatbot’s responses are properly formatted and secure before being displayed on the website or in the application.

6.4 User Interaction Examples

To illustrate the chatbot’s functioning, here are some examples of typical interactions:

- **Question**: "What are the prerequisites for the robotics course?"
- **Response**: "The prerequisites for the robotics course include basic knowledge in programming and applied mathematics. You can find more information [here](http://example.com/robotics)."
- **Question**: "When does the fall semester start?"
- **Response**: "The fall semester starts on August 21. You can view the full academic calendar [here](http://example.com/calendar)."
- **Question**: "Are there any scholarships available for international students?"
- **Response**: "Yes, several scholarships are available for international students. You can find the complete list of scholarships and eligibility criteria [here](http://example.com/scholarships)."

6.5 Technical Challenges and Solutions

The chatbot development involved several technical challenges:

- **Multilingual support**: Ensuring the chatbot can understand and respond in multiple languages required adjustments to the language models and integration of multilingual data.
- **Response time optimization**: While GPT-3.5 Turbo is fast, optimizing response times to ensure a smooth interaction required adjustments in request and response management.

In conclusion, the chatbot enriches the user's question, transmits it to the API GPT-3.5 Turbo to obtain a response in Markdown format, and then converts this response to HTML for display on a web page or in an application. This process allows the chatbot to provide precise and well-formatted responses to user questions.

7 Testing

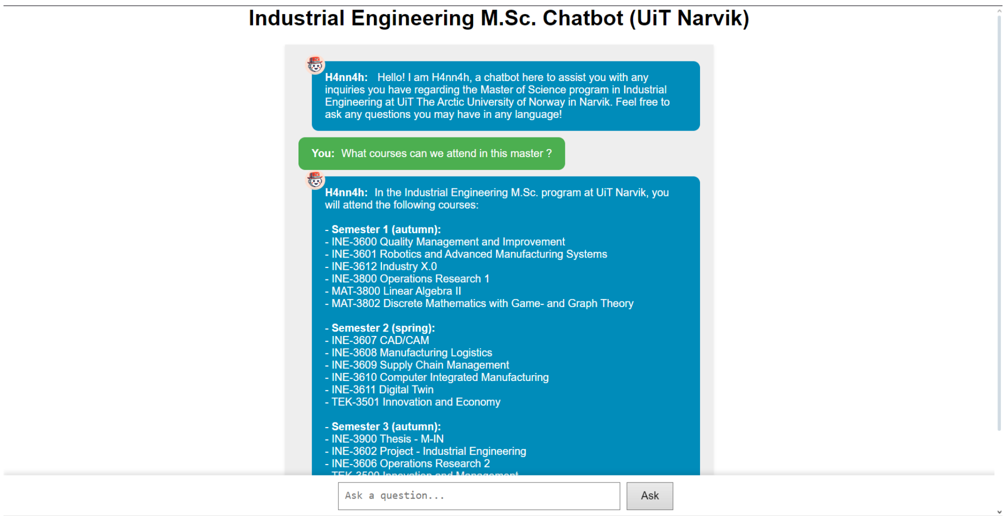

Figure 5: Chatbot Test Interface

The chatbot testing phase was conducted over a three-week period. A dedicated interface, containing only the chatbot, was created to allow for precise performance evaluation. The chatbot was deployed on a temporary server accessible to all participants, including UiT students and staff, to gather user feedback in real-world conditions.

The tests were designed to cover various aspects of user interaction with the chatbot. Users were encouraged to ask a variety of questions, ranging from general information about the Master's in Industrial Engineering program to specific details about courses and available services. Each interaction was followed by an evaluation, where users were asked to rate the chatbot's response on a scale of 1 to 5, with 1 representing an unsatisfactory response and 5 a very satisfactory response. They also had the option to add comments to help improve the system, providing valuable insights into the user experience and areas needing adjustments.

7.1 Test Methodology

The test methodology included several key steps:

1. **Participant selection**: Test participants included current students, professors, and administrative staff at UiT. This diversity allowed for a wide range of feedback.
2. **Data collection**: Interactions were recorded, and data was collected on the types of questions asked, responses provided, response time, and user evaluations.

3. **Data analysis**: Collected data was analyzed to evaluate the chatbot's performance, focusing on response accuracy, user satisfaction, and areas needing improvement.

During this testing phase, I collected several types of data: user questions, chatbot responses, user ratings, and optional user comments. Additionally, the average time the chatbot took to respond to each question was recorded.

7.2 Test Results

At the end of the three-week testing period, I was able to gather 170 responses. Of these 170 questions, 95% were rated between 4 and 5 out of 5, indicating high user satisfaction with the quality of the chatbot's responses. It should also be noted that questions were asked in different languages, including English, French, Norwegian, Sami, and Chinese. Although the sample of responses in languages other than English was relatively small, there does not seem to be a significant difference in response quality based on language.

Moreover, the average response time of the chatbot was measured at approximately 2.5 seconds. This quick response contributes to a smooth and satisfactory user experience.

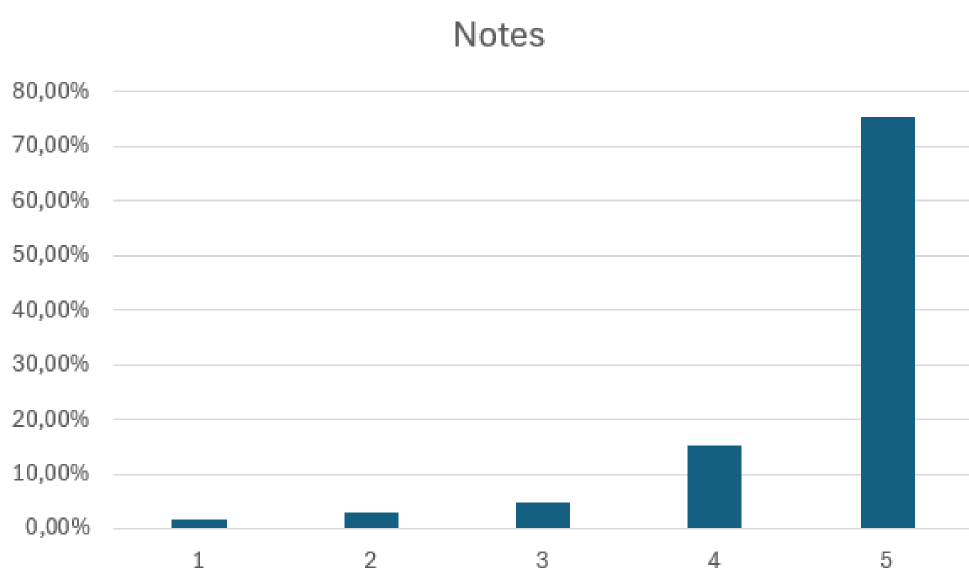

Figure 6: Distribution of User Ratings

7.3 Analysis of Results and User Feedback

The test results allowed for an in-depth analysis of the chatbot's performance and user perceptions. The distribution of user ratings, as illustrated in Figure 6, shows a predominance of high ratings, indicating general satisfaction with the quality of the provided responses.

User comments were particularly valuable in identifying potential areas for improvement. Several users expressed satisfaction with the chatbot's accuracy and relevance of responses, as well as its ability to handle questions in various languages. Here are some examples of user feedback:

- **Person A**: "The chatbot answered my question about the prerequisites for the robotics course very quickly and accurately. It's very useful to get an immediate response."
- **Person B**: "It lacks information on the research groups in the industrial engineering department."
- **Person C**: "It would be convenient to have precise links to the course web pages it mentions."

The comments also greatly helped complete the missing information in the chatbot and allowed for correcting any issues found during the testing phase. Here are some examples of improvements made following the feedback:

- **Enrichment of the knowledge base**: Added new responses for frequently asked questions not initially covered.
- **Improvement of the user interface**: Simplified the interface to make interactions more intuitive.

In conclusion, the tests validated the chatbot's effectiveness and reliability in its role as a virtual assistant for providing precise and contextualized information to UiT users. The insights gained from this testing phase are essential to guide future development efforts and ensure a continuous and improved user experience.

8 Integration into the Website

In anticipation of the chatbot's integration into the UiT website, I prepared an interface designed to provide easy and intuitive access to users. The popup will be positioned in the bottom right corner of the page, with a subtle animation to draw visitors' attention to the chatbot feature.

8.1 Popup Interface

The popup interface was carefully designed to adhere to the website's graphic charter. It is elegant and discreet, integrating the colors and visual style already present on the platform.

8.2 Key Features

When users access the chatbot popup, they are greeted by a pre-configured welcome message. This initial message serves to introduce the chatbot and explain its main purpose, which is to assist users in finding information about the Master's program and UiT.

8.3 Attention Animation

To effectively capture visitors' attention, the popup animation is designed to be subtle but attractive. It uses gentle pulses that naturally guide the eye toward the popup without disrupting the browsing experience.

8.4 Integration Objectives

The integration of the chatbot aims to improve information accessibility for users interested in the Master's program and UiT. By providing easy and quick access to responses, the chatbot contributes to a more satisfying and efficient user experience on our website.

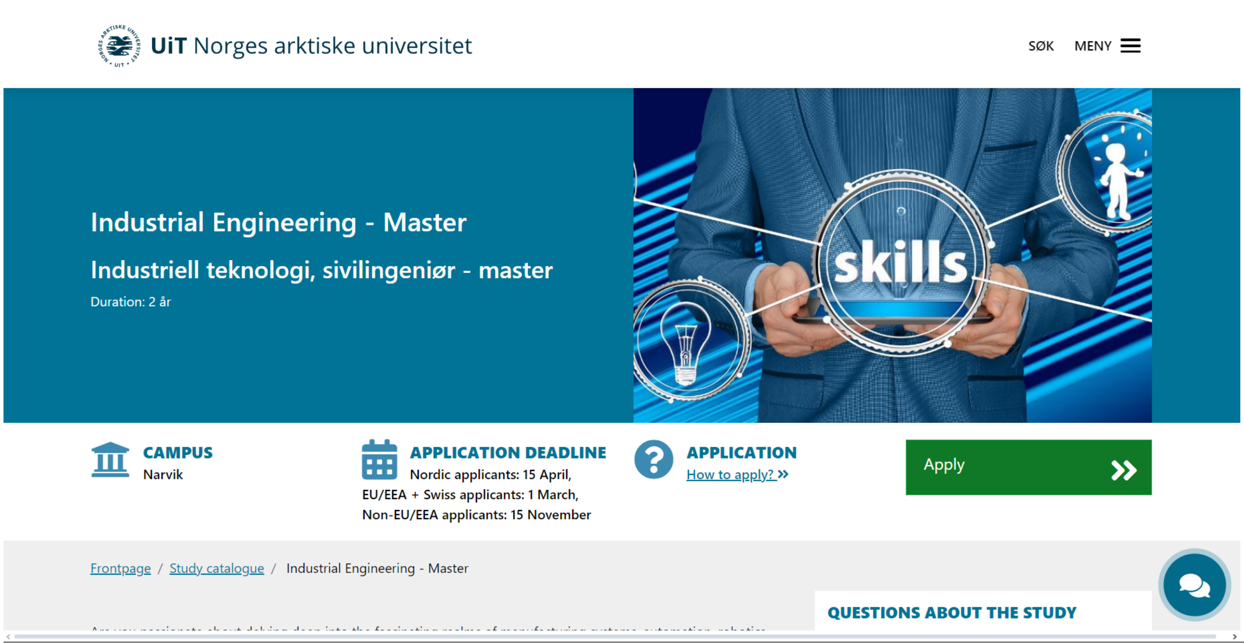

Figure 7: UiT Chatbot Interface Closed

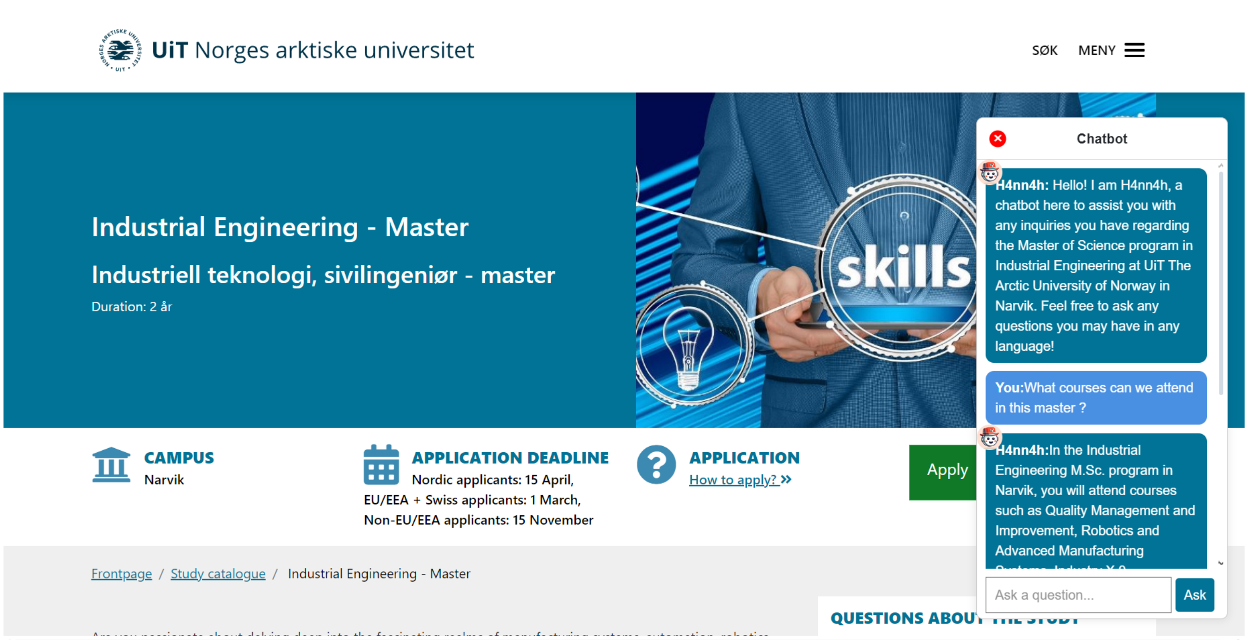

Figure 8: UiT Chatbot Interface Open

9 Simulation

The simulation phase is crucial to evaluate the chatbot's effectiveness and compare the results to those obtained two years ago. This section explores in detail the simulation environment used, the various scenarios tested, and the results obtained, with a particular focus on comparative analysis with previous simulations. Many of the values used are taken from simulations conducted two years ago in the document "Measuring the effectiveness of AI-enabled chatbots in customer service using AnyLogic"[3].

9.1 Simulation Environment Description

The simulations were conducted using AnyLogic, a discrete event modeling software known for its flexibility and ability to model complex systems to solve a wide variety of problems in many fields, for example, printing systems in a university [8], urban postal counter locations [9], and lean adoption in a factory [10]. The simulation environment simulates the flow of users arriving in a customer service system where each user first interacts with a chatbot. The chatbot is designed to handle an infinitely large number of users simultaneously, and its ability to satisfy users depends on its assigned intelligence.

If a user is not satisfied with the chatbot's response, they are redirected to a human agent. The human system is equipped with 1 to 2 agents capable of handling as many users as agents simultaneously. If a user waits too long without a response, they leave the process according to a wait time distribution defined by the Triangular(2min, 15min, 8min) function.

9.2 Simulation System Parameters

The key parameters of the simulation system are summarized below:

Parameter	**Value**
User arrival rate (min)	Triangular(0, 3, 1)
Chatbot intelligence (satisfaction percentage)	50.00%, 79.50%, 85.00%, 90.00%, 95.60%
Number of human agents	1 or 2
Chatbot response time (min)	Triangular(1, 10, 5)
Human agent response time (min)	Triangular(2, 12, 8)

Table 2: Simulation System Parameters

9.3 Simulation Scenarios and Results

I tested several simulation scenarios by varying the chatbot's intelligence. For the first scenario, I chose a 0% intelligence for the chatbot to represent a system without a chatbot, for the second I used the highest intelligence tested in the document "Measuring the effectiveness of AI-enabled chatbots in customer service using AnyLogic"[3], which is 50%. Then, I based the data collected during the actual tests: 79.5% corresponding to the percentage of responses rated 5/5, and 95.6% corresponding to the percentage of responses rated 4/5 or 5/5. I also conducted simulations with intermediate intelligences.

	Number of Human Agents	**Chatbot Intelligence**
1	5	0.00%
2	2	50.00%
3	2	79.50%
4	2	85.00%
5	2	90.00%
6	2	95.60%
7	1	95.60%

Table 3: Simulation System Parameters

	Human Utilization	**Total Users**	**Answered by Chatbot**	**Answered by Human**	**Departures**	**Average Time (s)**
1	99.00%	49 406	0.00%	39.42%	60.58%	618.51
2	100.00%	50 025	49.88%	15.61%	34.48%	614.54
3	92.00%	49 576	79.55%	14.54%	5.91%	444.13
4	80.00%	49 584	85.10%	12.59%	2.31%	405.18
5	59.00%	49 866	90.07%	9.40%	0.53%	371.41
6	27.00%	49 630	95.63%	4.34%	0.03%	340.20
7	48.00%	49 513	95.61%	3.84%	0.55%	341.15

Table 4: Results of Different Simulation Scenarios

9.4 Analysis of Simulation Results

The analysis of simulation results aims to evaluate the chatbot's effectiveness in the customer service system by comparing different scenarios with varying levels of intelligence and examining the impact compared to a system without a chatbot.

9.4.1 Comparison with a System without a Chatbot

In the first scenario where the chatbot had 0% intelligence, the results show an almost total reliance on human agents to respond to users:

- **Human Utilization:** 99.00%
- **Percentage of Responses by Human:** 39.42%
- **Percentage of Responses by Chatbot:** 0.00%
- **Departures (Abandonments):** 60.58%
- **Average Response Time:** 618.51 seconds

This demonstrates that without a chatbot, the workload on human agents is extremely high, with significantly longer average response times. Additionally, the high departure rate suggests high user dissatisfaction due to prolonged wait times.

9.4.2 Impact of the Chatbot (50.00% Intelligence)

In the second scenario with a chatbot having 50.00% intelligence, the results show a notable reduction in human workload:

- **Human Utilization:** 100.00%
- **Percentage of Responses by Chatbot:** 49.88%
- **Percentage of Responses by Human:** 15.61%
- **Departures (Abandonments):** 34.48%
- **Average Response Time:** 614.54 seconds

While the majority of responses are still handled by human agents, the introduction of the chatbot has reduced the overall average response time and distributed the workload more evenly. The departure rate has also decreased, indicating an improved user experience due to shorter wait times.

9.4.3 Increasing Chatbot Intelligence

As the chatbot's intelligence increases, there is a significant reduction in human workload and corresponding improvement in overall system performance:

- For 79.50% intelligence, the percentage of responses by the chatbot increases to 79.55%, further reducing human workload and response times.
- At 95.60% intelligence, the chatbot answers 95.63% of requests, leaving only 4.34% of responses to the human agent. The average response time is also considerably reduced to 340.20 seconds.

9.4.4 Reduction of Human Workload

Overall, the introduction and improvement of chatbot intelligence have a direct impact on reducing human workload:

- High percentages of responses by the chatbot in high intelligence scenarios (>85.00%) show that the chatbot can effectively handle the majority of interactions, allowing human agents to focus on more complex cases.
- This translates to shorter average response times and better allocation of available human resources, improving operational efficiency and customer satisfaction.

9.4.5 Reduction of Required Human Agents

The progressive reduction in the percentage of responses requiring human intervention suggests a potential decrease in the number of human agents needed to maintain effective service:

- Compared to the scenario without a chatbot, where practically all responses were handled by human agents, scenarios with higher chatbot intelligences require less human intervention.
- This could lead to operational cost savings and more efficient allocation of human resources, while maintaining high levels of customer satisfaction through faster and more accurate responses.

In conclusion, integrating a chatbot with appropriate intelligence not only improves customer service efficiency but also reduces reliance on human agents and optimizes the use of available resources. The detailed simulation results highlight the crucial importance of artificial intelligence in enhancing customer support processes, particularly in high-demand environments where efficient interaction management can make a significant difference.

10 Conclusion

This report describes the development and integration of a chatbot for the Master's in Industrial Engineering program at the University of Narvik (UiT) in Norway. The primary objective was to provide an automated solution to answer prospective students' questions, thereby improving information accessibility and reducing the workload of human agents. This project is part of a continuous innovation context in natural language processing (NLP) and chatbot technologies, with a particular focus on user satisfaction and operational efficiency.

The work carried out resulted in the design, development, implementation, and validation of a chatbot using the API GPT-3.5 Turbo, a large language model (LLM). The tests demonstrated high user satisfaction, with over 95% of responses rated between 4 and 5, and an average response time of 2.5 seconds. Additionally, simulations conducted with AnyLogic confirmed the chatbot's effectiveness, showing a notable reduction in human workload and abandonments compared to previous solutions.

However, the project is not entirely finished. Although the main objectives have been achieved, several aspects still require improvement. For example, continuously enriching the chatbot's knowledge base is essential to maintain the relevance and accuracy of the provided responses. Moreover, the full integration of the chatbot on the UiT website and end-user training are crucial steps to finalize.

The main difficulties encountered during the project were technical, particularly concerning training and adjusting the GPT-3.5 Turbo model to ensure precise and contextual responses. These challenges were resolved through an iterative approach and rigorous testing. Additionally, organizational challenges related to coordinating with various UiT stakeholders were overcome through effective communication and strategic planning.

The interest of this study lies in the concrete demonstration of LLMs' effectiveness in improving information services in an academic context. This project allowed for acquiring valuable skills in chatbot development, project management, and data analysis, while highlighting the importance of a user-centered approach for developing technological solutions.

Looking to the future, several prospects open up for this project. If the study is not entirely completed, there remains the task of fully integrating the chatbot on the UiT website and ensuring adequate end-user training. Additionally, the chatbot's knowledge base will need regular updates to reflect the most recent information. If the study is considered completed, the chatbot will be fully integrated into the UiT information system, serving as the main contact point for prospective students' questions. Continuous maintenance and regular updates will also be necessary to ensure its long-term effectiveness.

In conclusion, this project represents a significant advancement in using NLP and chatbot technologies to improve information access in the education field. The results obtained show a promising path for the future, with potential applications in other academic and institutional contexts.

References

[1] J. WEIZENBAUM, Eliza—a computer program for the study of natural language communication between man and machine, *Communications of the ACM*, vol. 9, no. 1, p. 36–45, 1966.

[2] T. B. BROWN, B. MANN, N. RYDER, M. SUBBIAH, J. KAPLAN, P. DHARIWAL, A. NEELAKANTAN, P. SHYAM, G. SASTRY, A. ASKELL, S. AGARWAL, A. HERBERT-VOSS, G. KRUEGER, T. HENIGHAN, R. CHILD, A. RAMESH, D. M. ZIEGLER, J. WU, C. WINTER, C. HESSE, M. CHEN, E. SIGLER, M. LITWIN, S. GRAY, B. CHESS, J. CLARK, C. BERNER, S. MCCANDLISH, A. RADFORD, I. SUTSKEVER et D. AMODEI, Language models are few-shot learners, 2020.

[3] X. SUN, H. YU et W. DENG SOLVANG, Measuring the effectiveness of ai-enabled chatbots in customer service using anylogic simulation. Manuscript in prepa ration, 2024.

[4] UIT THE ARCTIC UNIVERSITY OF NORWAY, Uit the arctic university of norway website. https://uit.no/startsida, 2024. Accessed: 2024-06-20.

[5] D. HAKKANI-TÜR, G. TUR, A. CELIKYILMAZ, Y.-N. CHEN, J. GAO, L. DENG et Y.-Y. WANG, Multi-Domain Joint Semantic Frame Parsing Using Bi-Directional RNN-LSTM, *in Proc. Interspeech 2016*, p. 715–719, 2016.

[6] H. CHEN, X. LIU, D. YIN et J. TANG, A survey on dialogue systems: Recent advances and new frontiers, *ACM SIGKDD Explorations Newsletter*, vol. 19, p. 2535, nov. 2017.

[7] J. DEVLIN, M.-W. CHANG, K. LEE et K. TOUTANOVA, BERT: Pre-training of deep bidirectional transformers for language understanding, *in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (J. BURSTEIN, C. DORAN et T. SOLORIO, éds), (Minneapolis, Minnesota), p. 4171–4186, Association for Computational Linguistics, juin 2019.

[8] X. SUN, H. YU et W. D. SOLVANG, Solving the location problem of printers in a university campus using p-median location model and anylogic simulation, *in Lecture Notes in Electrical Engineering*, vol. 634, p. 577–584, 2020.

[9] H. YU, X. SUN, W. D. SOLVANG et G. LAPORTE, Solving a real-world urban postal service network redesign problem, *Scientific Programming*, vol. 2021, p. 2021, 2021.

[10] A. AHMED, J. PAGE et J. OLSEN, Adopting lean six sigma to anylogic simulation in a manufacturing environment, *in 21st International Congress on Modelling and Simulation*, (Gold Coast, Australia), p. 29–35, 2015.