



**AFRICAN CENTRE OF EXCELLENCE IN DATA SCIENCE**  
**COLLEGE OF BUSINESS & ECONOMICS**

**A Multi-Lingual Conversational AI Chatbot for Effective Educational  
Consultations: A Case Study of ACE-DS, University of Rwanda**

**By**

**Kefas Rimannuskeb Galadima**

**Registration Number: 221031339**

**A dissertation submitted in partial fulfilment of the requirements for the degree of Master of  
Data Science in Data Mining**

**University of Rwanda, College of Business and Economics**

**Supervisor: Dr. Kizito Nkurikiyeyezu**

**July, 2023**

## Declaration

I declare that this dissertation entitled “**A Multi-Lingual Conversational AI Chatbot for Effective Educational Consultations: A Case Study of ACE-DS, University of Rwanda**” is the result of my own work and has not been submitted for any other degree at the University of Rwanda or any other institution.

**Names:** Kefas Rimamnuskeb Galadima

**Signature:** 

## Approval Sheet

This dissertation entitled “**A Multi-Lingual Conversational AI Chatbot for Effective Educational Consultations: A Case Study of ACE-DS, University of Rwanda**” written, and submitted by **Kefas Rimamnuskeb Galadima** in partial fulfilment of the requirements for the degree of Master of Science in Data Science majoring in Data Mining is hereby accepted and approved. The rate of plagiarism tested using Turnitin is **18%** which is less than **20%** accepted by the African Center of Excellence in Data Science (ACE-DS).



**Dr. Kizito Nkurikiyeyezu**  
*Supervisor*

---

*Head of Training*

## **Dedication**

This dissertation is dedicated to the memory of my late father, Mr. Kefas Galadima Utainriba, for believing in me and providing me with the unwavering support, encouraging me to explore and harness my hidden potentials. His belief in my abilities over the years have been the driving force behind my academic pursuits and success.

## **Acknowledgements**

I express my sincere gratitude to the Almighty God for his divine love, favour, and infinite mercy upon my life throughout my stay and Master's study at the ACE-DS, University of Rwanda.

My sincere appreciation goes to my dissertation supervisor Dr Kizito Nkurikiyeyezu, for his time and relentless support in accomplishing this research successfully. His advice, ideas, and review of my manuscript at different stages of this study have significantly impacted my research skills and contributed to the successful implementation of this work.

I also want to acknowledge and thank all my esteemed lecturers and the dedicated management staff at ACE-DS, for their firm support and contributions which have been instrumental in building and enhancing my intellectual, technical capacity and skills. I am indebted to them all for their kind support.

A special acknowledgement goes to the University of Rwanda via ACE-DS for finding me worthy for its master's programs and providing me with the funding (Scholarship) to pursue a master's degree in Data Science with a specialization in Data Mining. I am grateful for the trust and support extended to me through this sponsorship.

I owe a great deal of gratitude to my mother, Mrs Martha Kefas Galadima, my senior brother, Hon. Rima Shawulu Kwewum, my uncle, Mr Ayuba Galadima, and my siblings for their prayers, financial and emotional support throughout my academic career and endeavors thus far.

Finally, I want to sincerely thank all of my friends and colleagues, all the cohort 4 masters students at ACE-DS for their constant support, hospitality and encouragement. I would especially like to thank my friends whose constructive comments, criticism and recommendations helped accomplished this work successfully. I am eternally grateful for your love and support.

Thank you all.

***Kefas R. Galadima***

## ABSTRACT

*The requests for real-time simultaneous direct client-staff interactions and consultations in most industries and organizations is on the increase, particularly in the educational sector. These concurrent and numerous requests have led to queues, turmoil, and service delays, especially during and after working hours due to the shortage of employees and language barriers to cater for the non-stop services needed by clients. Artificial intelligence (AI), Natural Language Processing (NLP) and linguistics technologies have become a potent tool for creating innovative solutions to tackle these challenges. To this end, this study aims to better understand question-answering systems, with an emphasis on creating a multilingual conversational AI chatbot that will ultimately offer a workable solution in an educational setting.*

*The multi-lingual chatbot system developed in this research is designed to answer various consultation queries related to ACE-DS services. This covers tuition fees, admission procedures, postgraduate programs, facilities, program application requirements, program modules, travelling requirements for international students, student accommodations and so on. The study used annotated Frequently Asked Questions (FAQs) and corresponding answers about ACE-DS, incorporating sentiment and conversation data to train a Deep Learning (GRU RNN) algorithm. This enables the chatbot to engage in meaningful conversations, detect user moods, and provide appropriate responses based on its prediction. Additionally, the chatbot system incorporates language detection and translation capabilities, allowing users to engage the system in multi-lingual conversation.*

*The results of the system evaluation and performance based on analysis of the training data achieved an accuracy of 98%, with average weighted precision, recall and F1-score of 98%. Whereas, the testing evaluation yielded an accuracy of 99%, with an average weighted precision, recall and F1-score of 99%. The user satisfaction survey indicated that over 50% of the correspondents that participated in assessing the chatbot solution are highly satisfy with the performance of the system in correctness of grammar usage, contribution to a more efficient and time saving consultations, efficient use of preferred language, responses provided, prompt and correct response provisioning, and time saving in information acquisitions. The real-time testing of the chatbot was conducted locally on PC using the Visual Studio Code Editor running on a local server, and also through the internet with the assistance of NGROK.*

**Keywords:** Deep learning; Multilingual conversation; Chatbot; Artificial Intelligence;

# TABLE OF CONTENTS

<b>Declaration</b> .....	ii
<b>Approval Sheet</b> .....	iii
<b>Dedication</b> .....	iv
<b>ABSTRACT</b> .....	vi
<b>LIST OF ACRONYMS</b> .....	x
<b>CHAPTER ONE: INTRODUCTION</b> .....	1
<b>1.1 Background of the Study</b> .....	1
<b>1.2 Statement of the Problem</b> .....	3
<b>1.3 Aims and Objectives of the Study</b> .....	4
<b>1.4 Significance of the Study</b> .....	5
<b>1.5 Scope of the Study</b> .....	5
<b>CHAPTER TWO: LITERATURE REVIEW</b> .....	7
<b>2.1 Introduction</b> .....	7
<b>2.2 Artificial Intelligence as Conversational Tools</b> .....	7
<b>2.3 Deep Learning Algorithms</b> .....	9
<b>2.3.1 Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN)</b> .....	10
<b>2.3.2 Transformer Model</b> .....	12
<b>2.3.2.1 Language Models</b> .....	13
<b>2.3.3 Language translators and models in Generative AI</b> .....	13
<b>2.3.4 Major Challenges in Conversational AI</b> .....	15
<b>2.4 Research Gaps</b> .....	16
<b>2.5 Research Contribution</b> .....	17
<b>CHAPTER THREE: RESEARCH DESIGN AND METHODOLOGY</b> .....	18
<b>3.1 Introduction</b> .....	18
<b>3.2.1 The Agile Methodology</b> .....	19
<b>3.2.2 Research Design– Action Research</b> .....	22
<b>3.3 Modeling</b> .....	28
<b>3.3.1 Testing Techniques Deployed to Determine the Validity and Reliability</b> .....	30
<b>3.4 Evaluation Metrics</b> .....	31
<b>3.4.1 Evaluation of the Dialogue Model</b> .....	31
<b>3.5 Deployment</b> .....	33
<b>3.6 Development and Learning Tools</b> .....	35
<b>CHAPTER FOUR: SOFTWARE TESTING, RESULTS, EVALUATION AND DISCUSSION</b> .....	36
<b>4.1 Introduction</b> .....	36
<b>4.2 Test Analysis</b> .....	36
<b>4.2.1 Unit Testing</b> .....	36
<b>4.2.2 Integration Testing</b> .....	38

4.2.3	High-level Testing.....	39
4.3	Result Discussions and Evaluation.....	40
4.3.1	GUI, and User’s Queries/Response Documentation (Conversing with the Chatbot:)	40
4.3.2	Model Evaluation on the Training Data.....	42
4.3.3	Model Evaluation Using Validation Data.....	45
4.3.4	Model Evaluation Through Real Time Testing .....	47
4.3.5	User Satisfaction Evaluation Through Users Testing .....	49
4.3.6	Users Recommendation.....	50
<b>CHAPTER FIVE: CONCLUSION AND RECOMMENDATIONS .....</b>		<b>51</b>
5.1	Conclusion.....	51
5.2	Recommendations.....	52
<b>References: .....</b>		<b>54</b>

## List of Figures

Figure 3.1: Agile Process flow (Gaurav & Pradeep, 2012) .....	20
Figure 3.2: Project Plan/Management Using Trello Board.....	21
Figure 3.3: CRISP-DM Diagram (Data Science Process Alliance, 2023).....	22
Figure 3.4: Chatbot Dataset Snapshot.....	25
Figure 3.5: Data Preparation Diagram.....	26
Figure 3.6: Modeling Diagram .....	29
Figure 3.7: The Multilingual Chatbot GUI Presentation.....	30
Figure 3.8: Confusion Matrix (Kohavi and Provost, 1998).....	33
Figure 3.9: The Implementation of the Multi-lingual AI Chatbot .....	34
Figure 4.1: Training Loss of the Chatbot Model .....	38
Figure 4.2: Chatbot Interface with Chosen English Response language .....	41
Figure 4.3: Cross language Conversation with the chatbot system .....	42
Figure 4.4: Confusion Matrix of the model on the training data.....	44

**List of Tables and Listings**

Table 4.1: Classification Report of the Model on the Chatbot Training Data.....42

Table 4.2: Chatbot Model’s Classification Reports on Validation Data .....45

Table 4.3: Real Time Chatbot Testing/Evaluation.....48

Table 4.4: Users Satisfaction Experience with the research solution.....49

**LIST OF ACRONYMS**

ACE-DS	African Center of Excellence in Data Science
AI	Artificial Intelligence
ANN	Artificial Neural Network
APIs	Application Programming Interface
CNN	Convolutional Neural Network
CRISP-DM	Cross Industry Standard Process for Data Mining
CSS	Cascading Style Sheets
FAQs	Frequently Asked Questions
FN	False Negative
FP	False Positive
GRUs	Gated Recurrent Units
GUI	Graphical User Interface
HTML	Hypertext Markup Language
LLM	Large language models
LSTM	Long Short-Term Memory
ML	Machine Learning
MLP	Multilayered Perceptron
NHL	Natural Human Languages
NLG	Natural language generation
NLP	Natural Language Processing
NLT	Natural Language tool kits
NLU	Natural Language Understanding
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
TN	True Negative,
TP	True Positive
UI	User Interface
UR	University of Rwanda

## CHAPTER ONE: INTRODUCTION

### 1.1 Background of the Study

The advancement of Artificial Intelligence (AI) with a focus on Natural Language Processing (NLP) and Natural Language Understanding (NLU) is making significant strides in the areas of multilingual communication aims at bridging communication gaps across the globe e.g. the use of chatbots by organisations to replace human chat agents as a means of customer service agent. A conversational chatbot is a piece of intelligent computer software that mimics human-to-human communication over the Internet. These chatbots are integrated into messaging apps, mobile apps, or websites and are made to converse with users in natural human language (Lin et al., 2023). As noted in a study by Adam et al. (2021), conversational AI systems(chatbots), are now used to build intelligent conversational systems that are quickly replacing human chat and service agents. A study by Biswas (2020), further deduced that chatbot adoption is growing significantly across a number of sectors, including the dating, banking, and healthcare industries.

These chatbot systems are powered by AI and also have NLU capabilities. According to Biswas (2020), conversational AI systems have been predicted to be the "Next Interaction Layer," suggesting that conversational AI chatbots would soon replace the current interactions through websites and software apps. AI chatbot deployment primarily aims to improve non-stop service delivery by enterprises or organizations, ensuring that clients are attended to effectively, concurrently, and without tiredness day and night, removing queueing in their services.

Conversational AI with multiple language capabilities is an AI chatbot or conversational AI system that combines computational intelligence, linguistics, and computer science in its capabilities. It acts as an artificial agent and can process, analyze, understand, detect, and translate Natural Human Languages (NHL), as well as predict user queries using machine learning (ML) intelligence. A chatbot or bot is an AI system that can communicate intelligently by interacting with humans; In the case of voice-based enquiries, this might either be a spoken dialogue or a text message. According to Tarun et al. (2018), a conversational agent converses with people in a particular industry or on a particular topic using input in natural language sentences. The main purpose of chatbots is information provisioning, although most of the time it is accessed online, it may be used with local computers and mobile devices as well. While some chatbots also offer multilingual functionality, most chatbots are created with the English language

in mind. Google Translate, for example, supports 103 languages, which is more than any other multilingual agent (Mazhar, 2020).

Yang et al. (2023) in their study built a multilingual NLP-based AI chatbot as part of their research during the COVID-19 pandemic which can accurately address open-ended questions about COVID-19. This Chatbot was used to support healthcare delivery and pandemic preparedness. Thus, chatbots may be essential in situations where there are few opportunities for face-to-face interactions with medical professionals or there is a work overload as this can be done in an interactive way similar to the way patients and doctors typically communicate in exchanging health-related information. Voice chatbots can automate acute care triaging, remote monitoring, and chronic disease management. A study by Anand et al. (2023) noted that by providing users with speedy services, chatbots have recently changed the dynamics of customer care services. That is, Customers' queries or consultations can be addressed by conversational AI chatbots efficiently and in real time. Taking these into consideration, this study suggests that chatbots may play a crucial role in the easy distribution and accessibility of reliable information across a variety of fields, specifically in the educational setting that serves as the backdrop for this research.

The potential societal benefits of using chatbots and conversational user interfaces, as well as how this might improve self-help groups, are discussed in this study because Chatbots have the potential to easily increase the accessibility, affordability, and availability of key services in industries. In addition, this research suggests the design and implementation of a multilingual conversational AI system, also known as a Multi-lingual bot, that will be equipped with NLP, AI and NLU capabilities to detect, understand, and translate user input messages in multiple languages (e.g., English, French, Kinyarwanda, Swahili, Hausa, Igbo etc.). The system will have the capability to recognize and answer users' questions and requests in a multilingual dialogue. In educational institutions, individuals frequently request for information on scholarships, fees, admission details and applications, travelling guidelines for international students, questions about college programs, filing complaints, and requests for help, among others. To deliver more individualized and pertinent responses, the system will be built to anticipate various user inquiries and demands and provide appropriate responses. In order to facilitate cross-language communication, this will entail the creation of an NLP model that is tailored to a particular popular language (English) as well as the incorporation of language detection and translation technologies to handle the aspect of cross-language conversation.

## **1.2 Statement of the Problem**

There is often an overwhelming number of requests for first-hand interaction or consultation between clients and staff of organizations/establishment. These requests often happen concurrently resulting in commotions, queuing, and delays in the rendering of services between clients and their servers or the organizations. Sometimes it takes hours to complete a single client's requests, forcing all other clients to wait in line for the necessary execution time. This manual method of consultation is tedious, time demanding, inefficient, characterized by poor record keeping/storage, ineffective (i.e., the limited language of interaction, e.g., English or Kinyarwanda as the case may be), subject to poor performance due to the imbalanced client-to-staff ratio fostered by fatigue; and unsatisfactory to the overall client population. The solution to be provided in this study is aimed at the African Center of Excellence in Data Science (ACE-DS), University of Rwanda (UR) which is plagued with the aforementioned challenges. The ACE-DS is an academic institution where the general public, prospective students or international students in particular find it challenging in making consultations or acquiring suitable information regarding studying at ACE-DS.

The chatbot solution for this study only addresses educational queries such as About ACEDS, admission requirements, accommodation, tuition fees, the profiles of top management staff, available courses and requirements, postgraduates' programs, admission process, funding, student accommodation, transportation system around ACE-DS and in Kigali, why choose ACE-DS as a study destination, travel and Visa requirements for international students and so on associated with ACE-DS. The chatbot does not cover queries that are personal e.g. inquiries of staff marital and religious orientation, age and residence; inquiries on the architecture and logistics of the ACE-DS for security and privacy concerns. Furthermore, these queries can be traversed in 1 - 100 languages (English Swahili and Kinyarwanda inclusive) which makes it equally beneficial in the sense that users can converse with the system in a range of languages in which they are much more comfortable and fluent in.

The multilingual conversational AI system (Chatbot) is the ideal solution to adopt because it is essentially an automated system that simulates a human consultant/member of an organization that serves multiple clients in cross-language communication. This system uses a Machine learning (ML) algorithm to understand client queries/intents and can attend to a large number of users simultaneously and in multiple languages thereby enhancing efficiency and client satisfaction.

### **1.3 Aims and Objectives of the Study**

The aim of this study is to investigate and implement a Conversational AI System that can process and answer users' consultations queries. This system will act as an information provisioning agent in an educational institution (ACEDS) and have the ability to effectively process queries in multiple languages in order to assist the general public like students, prospective students, and visitors.

The objectives of this research include:

- i. An evaluation of manual systems of consultation and an analysis of contemporary AI chatbots to determine the benefits and drawbacks of their various features, and then the development of the intended chatbot system using the concepts and methods that were found useful.
- ii. To create an AI-powered chatbot with a user-friendly interface which can be measured by its responsiveness; clarity, conciseness and non-ambiguity of responses provided; easy navigation on the menu, the use of simple design and non-conflicting colors; and efficient/speedy execution of queries/responses well-formatted to the user interface. Consequently, the chatbot will enable any number of users to connect to a server and securely carry out consultation concurrently. The metric for verifying this concurrency are– the number, time and duration of queries received by the server; and the accuracy values retrieved for each of these queries.
- iii. To study and integrate Language detector and translator APIs for detecting and translating users' queries and responses to their preferred language. The Language detectors and translators e.g. Google translate, DeepL translate and EasyNMT translator, and language detection APIs will be accessed via open source/free resources. These APIs will be studied via the documentation and tutorials appended to them and furthermore firsthand experimentation will be carried out on them by the researcher by deploying them locally on datasets and further integrating with the study solution.
- iv. To build a system that can analyze, understand, and automatically detect the language in which a user is conversing and correctly respond to users' queries in cross-language communication.
- v. To assess the effectiveness of the system by evaluating the machine learning model using metrics such as classification reports and confusion matrix detailing the F1 score, accuracy, precision, and the recall of the model. Assess the veracity and reliability of the information the chatbot provides via the unit, high level, users, and real-time testing of the system, noting the confidence level and accuracy of its predictions, and the degree to which users are generally satisfied with the system represented by satisfaction rating.

## **1.4 Significance of the Study**

The AI-driven multi-lingual conversational system proposed in this study has the capability to learn from users' queries and predict what sort of responses to provide to users. This study will essentially help readers better understand the application of multilingual intelligent chatbots with the necessary tools and knowledge, and how it can be deployed in real-life settings. e.g. educational and health institutions can use multilingual intelligent bot to ensure that consultation and queries from clients or customers are attended to at all times (24/7) and their data are secured from intruders. The study has the following significance:

- i. The chatbot will be highly scalable as an arbitrary number of clients can be attended to simultaneously which eliminates the need for queuing thereby fostering efficiency in consultation and information provisioning services. This efficiency is achieved via an elimination of waiting time because all client queries are attended to on the go.
- ii. Clients or users can interact with the system at all times of the day(day/night) all through the week which is not the case in the manual system i.e. it is dependable and its availability is undebatable.
- iii. A key significance of this study is to develop a chatbot that will not only respond to queries but will do so empathically with art, hence the chatbot will not just be robotic but empathic. Similarly, according to Firdaus, Ekbal, and Bhattacharyya (2020), a highly functional chatbot system prioritizes and establishes a visceral connection with clients.
- iv. The chatbot system is not prone to exhaustion or fatigue, indicating that it will be reliable and effective.
- v. The chatbot system will offer services in a wide range of human languages as opposed to the conventional system where a staff member might only be proficient in 1 or 2 languages. Also, taking the case of Amzon Alexa and Google Assistant where the system only communicates with users in a single language (Zhou,2022). This ensures that convenience is facilitated on the clients' part because they can present their cases in the language they are most comfortable and fluent in.

## **1.5 Scope of the Study**

This study will provide a concise discussion on AI, the contemporary chatbot systems, and ultimately develops a Multi-lingual Conversational AI Chatbot system that enables clients to interact with a human-simulated consultant in 1 to over 50 languages only. The application to be developed is web-based, hence the scope of the study does not cover providing a mobile or standalone application. The chatbot solution in the context of this research will be deployed and experimented with in an educational institution only

i.e. ACE-DS at the University of Rwanda, it does not cover consultancy services in any other discipline like business, banking etc. Moreover, in the educational institution i.e. ACEDS, only the following specific functionalities will be provided for the purpose of this study: multilingual interactions, predictions of client queries or consultations with respect to context, deployment of deep learning algorithms (i.e. Gated Recurrent Units (GRUs) kind of Recurrent Neural Network) to learn and respond to queries thereby providing the functionalities of a virtual consultant. Further, evaluating the interactions between clients and the system by undertaking data analysis and data visualization (presenting insightful information derived via real-time testing, model training and validation of the chatbot solution).

The interactions and responses between the user/client and the solution will be conducted only in written form i.e. text/typing; other means of communication such as verbal/audio, and video/visual will not be implemented in the context of this research. The type of data to be used for this research is essentially qualitative data in only textual form. It excludes data presented in the form of charts, tables, graphs, audio, images and video. Albeit, it encompasses a mix of secondary and primary data only.

## **CHAPTER TWO: LITERATURE REVIEW**

### **2.1 Introduction**

The application of AI, NLU and NLP have a wide range of applications in industries, organizations and in our personal and social lives. Our daily and professional life is characterized by repetitive, cumbersome and exhaustive tasks; and in some cases, rather than foster mastery and excellence, the repetitive execution of such tasks only facilitates poor productivity and fatigue. AI has been adopted in personal life in the form of voice assistants in smartphones/gadgets, image recognition used for security, health and agriculture, and in maps for directions and navigations. In the aspect of professional setting, AI has been used in communications, e.g. auto-correct in text editors, search and recommendation algorithms, and determination of users' intents (NLU) as in chatbots, e-payments, medical diagnosis, and digital assistants, to mention a few.

Additionally, the application of AI with NLP and NLU is making great transformations in bridging the communication gaps across communities particularly, in the areas of communication and understanding of world Languages e.g. in sentiment analysis, language-to-language translations, text to voice, voice to text amongst others.

### **2.2 Artificial Intelligence as Conversational Tools**

Health sector is one of the essential areas in which AI has been deployed to serve as a conversational tool, e.g. in telemedicine. The deployment of AI in the health services is essential in some cases where consultation is imminent but physical contact between patients and clinicians is impeded by the presence of communicable diseases, hence, virtual interaction is necessary. Although communication via audio or video could be deployed but in the cases of a pandemic where the patients outnumber physicians or health facilities such as the recent occurrence of the covid-19 pandemic, a virtual assistant will be required to help attend to non-critical cases. According to Bharti et al. (2020), telemedicine has been instrumental in the midst of the recent covid-19 pandemic by enabling patients to obtain supportive care without the need of going to a hospital physically by using a conversational artificial intelligence-based health program. Hence, telehealth will drastically and quickly change in-person care to patient consultation through the internet via conversational AI systems. Furthermore, in order to improve patient quick access to healthcare information and to take advantage of the capabilities of AI to close the gap between access to healthcare providers, their work led to the development of a conversational chatbot system called "Aapka Chikitsak" on Google Cloud Platform (GCP) for telehealth provisioning in India. In order to provide timely care and high-quality treatment, this conversational application has

successfully lowered the obstacles to accessing healthcare-related services, and facilities, and secured intelligent consultations remotely.

One of the lapses in the virtual medical AI assistant as deduced from the work of Bharti et al.(2020) is the absence of empathy and a visceral connection between patients and the artificial agent. A crucial component of customer care systems is client satisfaction, and such systems must act courteously when responding to user requests/queries and demands (Firdaus, Ekbal, & Bhattacharyya, 2020). A study by Koehler et al.(2018) argues that the communication and feelings which a client experience while receiving a service is more important than the quality of the services provided and hence should have a higher priority. They emphasized that human satisfaction is essentially derived from the fulfilment of an emotion. Furthermore, they expounded that many solutions exist for artificial intelligence (AI) to improve service processes, but many initiatives merely advocate replacing employees with chatbot systems. Whereas, in their work, it was pointed out that three intelligent assistants that help service personnel with their complex responsibilities instead of pure automation that focuses on consumer self-service are often not customized to individual client needs (psychological and emotional needs). The solution comprises these components, the scribe, the skill manager, and also the background knowledge worker. As regards the need for empathy, the background knowledge worker is importantly responsible for acquiring information about clients to determine their behavioural patterns, longings and psyche. This component will help adapt the choice of words and communication process to appeal to the client's emotional expectations. Additionally, from the works of Firdaus, Ekbal, and Bhattacharyya (2020), one of the ways to resolve the problem of empathy is to accrue data from a highly interactive social networking platforms such as Twitter and Facebook, that constitutes all sorts of interactions, objective, sentimental, emotional, courteous and generic responses between users and/or customer care agents. These data are used to train deep learning machine learning frameworks that can simultaneously handle different interactions both formal, informal, and emotional, and incorporate empathy or polite behaviour in the customer care AI agent's responses.

In most contemporary AI-driven dialogue/chatbot systems, the number of languages is very limited, mostly in a range of 1-5. A study by Zhou (2022) affirmed that current dialogue systems, such as Amazon Alexa and Google Assistant, frequently only communicate with human users in a single language, which restricts the usage of these systems to conducting situated dialogues that require visual awareness or communicating with users who speak various languages. A particular barrier to the communication between these systems and users is that the mode of communication is limited to written and oral forms, thereby excluding other users such as the deaf and dumb who essentially use sign language. The work of Zhou (2022) also explored the possibility of creating a system that interacts using visual/sign, oral,

and written language. Consequently, the solution proposed by Zhou was to investigate ways to make it easier for an agent to link speech and vision to actions in a conversational setting. Furthermore, to train task-oriented visual conversation systems, it suggested an alternate learning method that combines supervised learning with reinforcement learning algorithms to achieve a better balance between dialogue response quality and policy efficacy.

According to Razumovskaia et al. (2021), one of the major challenges mitigating the advancement of chatbot systems is the unavailability of sufficient data for undertaking machine learning training, particularly in not-so-popular languages. Hence, data collection with regard to these languages is expensive and time-consuming and therefore, the majority of currently used methods for multilingual chat systems rely on cross-lingual transfer from languages with abundant resources to languages with fewer resources.

## **2.3 Deep Learning Algorithms**

Deep learning is a multi-layered processing ML algorithm based on the architecture of an artificial neural network (ANN) and has the capability to extract deep features from data. Computational models with many processing layers can learn data representations with different levels of abstraction thanks to deep learning ML algorithms. Deep learning employs the backpropagation algorithm to suggest how a machine's internal parameters adjust its weights and biases to obtain optimum results, which are used to derive the representation in each layer from the representation in the previous layer. In a nutshell, deep learning can uncover subtle structures in massive datasets. These techniques have considerably improved the state-of-the-art in a range of domains, including drug discovery and genomics, object identification, visual object recognition in computer vision, and speech recognition in NLP (LeCun, Bengio, and Hinton, 2015).

The work of Mehrish et al. (2023) further pointed out that with the advent of deep learning, the field of speech processing in NLP has undergone a transformative paradigm shift. Multiple processing layers in deep learning have allowed for the development of sophisticated models capable of extracting intricate features from speech data. This advancement has paved the way for unprecedented advancements in speech recognition, text-to-speech synthesis, automatic speech recognition, and emotion recognition, propelling the performance of these tasks to previously unheard-of levels. Deep learning techniques' power has paved the way for new avenues of research and innovation in the field of speech processing, with far-reaching implications for a variety of industries and applications.

Artificial intelligence's ability to critically classify text, and medical images, and learn the hidden underlying patterns in data depends on deep learning algorithms. Radiologists and orthopedic surgeons' lives are made simpler by deep learning algorithms, which give them quicker, more accurate results in real time than the conventional system. Nevertheless, the performance of the conventional deep learning algorithms has been reached (Narayan et al., 2023). The medical research community made substantial use of deep learning algorithms for the analysis of medical pictures and data during the pandemic. Deep-learning techniques, which are AI-based methodologies, have been widely employed (Aslani & Jacob, 2023). In this study, deep learning is crucial since its subdomains, including neural networks such as recurrent neural networks (RNN) and transformer models, are required in the implementation of NLP applications.

### **2.3.1 Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN)**

ANN are a collection of interconnected nodes or neurons, that process information in a manner similar to how the human brain does. NLP, computer vision, financial modelling, and predictive modelling are just a few of the areas where ANNs and CNNs have been successfully applied with substantial results. One of its most important characteristics is the ability of these algorithms to grow and learn from their experiences, backpropagation, a method that reduces errors between the output and the intended result by adjusting the weights of connections between neurons. This is used to achieve this learning, this technique has made ANNs a powerful tool for pattern recognition and classification in different fields (Kurani et al., 2023).

The properties and functions of the nervous system have long been defined by neuroscientists, and they are now getting closer to comprehending how the brain performs its functions. However, the question of "why" brains work the way they do isn't brought up as frequently. Thanks to the new ability to optimize ANNs for performance on human-like tasks, the research by Kanwisher, Khosla, and Dobs (2023) attempts to answer these "why" questions by figuring out when the qualities of networks optimized for a given task match the behavioural and neural characteristics of humans performing the same activity. Additionally, cutting-edge methodologies and technologies were widely applied in the study by Kurani et al. (2023), and ANN is used as a hybrid model, such as ANN-MLP and GARCH-MLP, which combines the backpropagation algorithm with Multilayer Feed-forward network to produce superior results.

CNN and RNNs, are two types of ANNs used in deep learning which have been increasingly popular in the fields of computer vision and natural language processing in recent years. While RNNs have been

used on sequence data to produce text, carry out language translation, and forecast stock prices, CNNs a computer vision algorithm has been demonstrated to be extremely effective for picture recognition (Sutskever, Vinyals, & Le, 2014).

ANNs have had several breakthroughs with exceptional results, however, there are still some barriers mitigating its performance that need to be handled in order to achieve better and more sophisticated results. The requirement for huge datasets to guarantee accuracy and avoid overfitting is one of the difficulties in ANNs. Another issue is that ANNs are not easily interpretable, which makes it challenging to comprehend how the ANN makes its decisions (Goodfellow, Bengio, & Courville, 2016).

Researchers have suggested a number of methods to solve these problems, such as employing transfer learning, which uses trained models to perform better on smaller datasets, and the attention mechanisms or transformer algorithm, which enable the network to concentrate or pay attention to particular portions of the data. Furthermore, explainable AI (XAI) research tries to make ANNs transparent and understandable by using visualization approaches and feature importance analyses (Ribeiro, Singh, & Guestrin, 2016).

ANNs have been a valuable tool in many fields, providing improved accuracy and efficiency in data processing and analysis. However, there are still challenges that need to be addressed in order to further improve the performance and interpretability of ANNs. Further research into advanced techniques and applications of ANNs will continue to expand their capabilities and potential impact in various industries.

According to Ghimire, Kil, and Kim (2022), a CNNs' ability to learn comes from a combination of several feature extraction layers that make full use of a lot of data. Although CNN frequently replaces conventional hand-engineered features in current systems, they frequently require significant processing power and memory resources. The study further showed that there are three areas which are paramount to increase the effectiveness of deep learning research which are: quantized/binarized models, efficient architectures, and resource-constrained systems. The following are some popular ANN techniques that are frequently employed in the creation of conversational chatbot systems:

- i. Recurrent neural networks (RNNs): These are a category of neural networks that can handle sequential data. RNNs are frequently employed in the implementation of chatbot systems because they feature a recurrent connection that enables them to keep an internal state or memory making them well-suited for handling conversation context.
- ii. Long Short-Term Memory (LSTM): LSTM is a particular kind of RNN deep learning architecture that rectifies the issue of vanishing points gradients in the model and is efficient at

identifying long-term dependencies in sequential data. Also, LSTM is well known for understanding and producing responses similar to that of humans, which makes it essential in chatbot systems implementation.

- iii. Gated Recurrent Units (GRUs): GRU is another RNN architecture that has LSTM features embedded in it but with a simpler structure. In addition to the LSTM features, GRUs have information flow-controlling gating techniques that make them computationally more efficient than LSTMs while still capturing the long-term dependencies in LSTM architecture.
- iv. Transformers: Transformers is a neural network design that is relatively new and popular for chatbots and other NLP applications. Transformers architecture employs a self-attention mechanism to identify connections between various words or tokens in a sentence and is used in a variety of NLP-related tasks. They have proven the best performances in their applications with the highest levels of performance and are proficient at handling context and producing coherent responses in NLP.
- v. Sequence-to-Sequence (Seq2Seq) Models: When it comes to machine translation and chatbot development, Seq2Seq models are mostly used because it has an encoder-decoder design, and RNNs, LSTMs, GRUs, or Transformers models can be used as the underlying architecture. In a chatbot system, the design processes the user's inquiry with the help of an encoder, and the chatbot's response is produced by the decoder.

### **2.3.2 Transformer Model**

The Transformer model, which has found success in many artificial intelligence fields, has shown great promise in modeling graph-structured data types. To date, a wide range of Transformers models have been proposed to adapt to graph-structured data (Min et al., 2022).

Transformer models can scale to billions of parameters, and model parallelism is necessary to distribute model parameters, activations, and optimizer state among devices, allowing them to fit into device memory and be trainable in a reasonable length of time. However, there are restrictions on scale model parallelism, even though it reduces the number of parameters per device linearly, for instance, when the parallel size is doubled. Splitting a model across numerous devices becomes ineffective due to tensor-level model parallelism's increased communication needs and smaller, less effective matrix multiplications (Korthikanti et al., 2023).

### **2.3.2.1 Language Models**

Transformers, a type of deep neural network algorithm designed specifically for processing sequential input, including text, have emerged as one of the most significant developments in AI language processing models. The transformer design of these AI models enables them to produce text that is both coherent and semantically intelligible, as well as to comprehend the contextual linkages between words in a sentence (Devlin et al., 2019). One of the most noteworthy applications of a transformer-based AI language model is the ChatGPT, which was created by OpenAI, and can carry out a number of language-related tasks, including text production, question answering, and even poetry composition. These can be accomplished using the highly developed ChatGPT language model, which was trained on large amounts of data (Brown et al., 2020). Industry experts have highly complimented ChatGPT's exceptional performance as "remarkable" and "impressive" (Li et al., 2021).

Unsupervised learning AI algorithms, which enables models to be trained on massive amounts of data without explicit supervision (Brown et al., 2020), is another significant advancement in AI language processing. Thanks to the capability of unsupervised learning algorithms, these models can perform tasks like text creation and question answering without the need for manual labelling (Liu et al., 2021). Since ChatGPT has been trained on a huge corpus of text data and can generate coherent and semantically relevant language text even without active supervision (Brown et al., 2020), this is especially important.

AI language processing models, like ChatGPT, have been put to the test against human performance on a variety of tasks ranging from question answering, text generation, and even poetry composition whereby ChatCPT was able to outshine human performance. The fact that these models routinely beat human performance serves as evidence of their extraordinary abilities (Brown et al., 2020). For instance, in a challenge that required question-answering, ChatGPT recently outperformed human performance in terms of accuracy and speed (Li et al., 2021), this shows how these models may automate a range of language-related tasks, offering important benefits in terms of speed, accuracy, reliability and scalability.

### **2.3.3 Language translators and models in Generative AI**

Generative AI models have proven to be efficient through their brilliant performance on a range of NLP tasks, including language production/translation, inference, and comprehension. However, the capabilities and constraints of these models are one of the most crucial concerns that the AI community is currently discussing, and it is obvious that evaluating generative AI is rather difficult. It is uncertain how well generative large language models (LLMs) can understand and produce other languages because the majority of research on these models is limited to English only.

Conversational AI models like OpenAI's ChatGPT and Google Bard AI serve as examples of how quickly AI and language processing have advanced over the time and its application to social and professional settings. A large amount of data has been used to train large language models named ChatGPT and Google Bard, which are now capable of a variety of language-related tasks like text generation, question answering, and even poetry composition with super-human performance. A wide range of applications have resulted from the extraordinary performance of AI, which has caught the interest of both academics, professionals and business executives. However, the use of conversational AI models like ChatGPT and Bard AI has brought about significant ethical and societal problems which several industries like the academic environment are very concerned about. Therefore, ethical issues and values must serve as the foundation for both the development and deployment of these large language models (Mattas, 2023). Hence, considering the widespread use of conversational AI models like Bard and ChatGPT, significant ethical and sociological issues must be made to curtail its negative impacts on societies and in the research community. Concerns have been raised about the possibility that AI systems could feed stereotypes and preconceptions as well as the possible impacts they might have on the labor market (Powles, 2020). Additionally, there are concerns about the effects of using AI systems to complete tasks that were previously completed by human personnel (Brouwer et al., 2021).

ChatGPT and Bard are examples of conversational AI models that have the capability to revolutionize social and professional life through how we engage with technology and how we communicate and acquire information. For instance, chatbots can interact with clients by responding to consumers' inquiries quickly and individually without the need for human intervention (Müller et al., 2019). Hence, using chatbots to inform patients about their health, medications, and other health-related issues, could have significant ramifications in sectors like healthcare (Chung et al., 2021).

Owing to the literature consulted thus far, and the first-hand experiments conducted by the researchers, it was further deduced that these Conversational AI models could communicate across many languages essentially because they have some language translators and detections API integrated into the system. Examples of some of these language translator's APIs are Google translate, DeepL translate and EasyNMT translator APIs. Some examples of language detection APIs are langdetect, text blob and langid APIs. These language translators and detectors are Python packages that could detect and decode written sentences in the form of text inputs from the source language to the user's target language with the objective to make the target language appear as though it is the original language of the written content when read by individuals using the translator. The implementation of language detections and translations is applicable to the context of this study for the purpose of multilingual conversation.

### **2.3.4 Major Challenges in Conversational AI**

NLU is one of the major obstacles in conversational AI systems, as NLU describes a computer system's capacity or ability to understand the meaning of spoken or written language; This necessitates that the system comprehends natural language's context, syntax, and semantics. Natural language, on the other hand, is complicated and fluctuates according to the speaker, the situation, and the culture. To enhance NLU, a number of methods have been developed, including rule-based strategies, statistical techniques, and deep learning approaches. These methods frequently fail to deal with ambiguity, sarcasm, and cultural quirks, which causes misunderstandings (Bharti et al., 2020).

From the research by Bharti et al. (2020), it was further inferred that Natural language generation (NLG) is another challenge for conversational AI. This is because NLG requires developing logical, convincing solutions to customer inquiries, i.e. the system must integrate knowledge, logic, and language creation skills in order to achieve this. Robotic and repeated responses are typically given by NLG systems, which decreases user satisfaction and interest. NLG techniques have been developed to address this challenge, including template-based tactics and neural language models. However, these techniques typically fail to generate appropriate ranges of responses that are suitable for the existing conditions.

Privacy and ethical issues are two more additional major challenges for conversational AI systems because sensitive personal data that conversational agents routinely collect includes users' locations, interests, and health information which most times are of critical concern to user's privacy. This raises concerns about the security and privacy of user data as well as potential data abuse by mostly the system developers or owners of the software. Additionally, conversational AI systems may promote preconceptions and stereotypes due to the data used to train them, as a result, some user groups, such as women, minorities, and people with disabilities, may experience discriminatory treatment. To address this, numerous initiatives and guidelines, such as the Fairness, Accountability, and Transparency in Machine Learning (FAT/ML) principle, have been released to address these issues (Biswas, 2020).

Conversational AI faces additional challenges in terms of interoperability and integration with other software systems in the IT industries. To provide accurate and relevant responses, conversational agents may need to be integrated with various data sources, such as customer databases and service logs to have the effectiveness to produce accurate and relevant responses. Hence, integrating conversational AI agents with existing systems, on the other hand, can be difficult and time consuming. Furthermore, conversational AI agents may need to interact with other conversational agents or virtual assistants, which necessitates standards for standardization and interoperability with such systems (Adam, Wessel, & Benlian, 2021).

In summary, NLU, NLG, privacy and ethical issues, interoperability, and integration are the major challenges in the deployment of conversational AI systems in solving real-world problems. These challenges necessitate interdisciplinary research and collaboration among computer scientists, linguists, psychologists, and ethicists. Hence, addressing these issues will allow for the creation of trustworthy, more effective and user-friendly conversational AI systems, which have the potential to revolutionize human-computer interaction.

## **2.4 Research Gaps**

From the available research conducted thus far, only little study was found on how a user's linguistic background affects their engagement with multilingual chatbots systems which can be inferred that there is no scientific evidence to support the notion that a user's level of fluency in a given language would influence how well they engage with a chatbot that uses that language. A greater understanding of how to include cultural and regional elements into multilingual conversational AI systems is also necessary to increase its accuracy, precision and performance. Without taking these elements into account, a chatbot's efficacy may be constrained because different locations and cultures have different communication preferences and conversational norms.

Another area of research for multilingual conversational AI systems is data privacy and security issues because these systems may interact with users who speak different languages and come from different countries, hence, safeguarding user data may create particular difficulties that call for more research. There is a large research deficit in this area despite the growing popularity of multilingual conversational AI systems. Future research in this area may focus on standardization, linguistic diversity, cultural and regional aspects, as well as data privacy and security issues which are indispensable in addressing these problems. These will make Multilingual AI systems to be more effective and have a greater potential impact on businesses and the larger customer base.

Consequently, the literature evaluated thus far revealed that in contemporary chatbot solutions, the range of languages available for interaction and consultation between the client and the system is typically in the range of 1 – 5 and also, it is often built based on dominant languages such as English, French, Spanish, Chinese, and German. Remote languages in Africa such as Kinyarwanda, Hausa, Igbo, Swahili, e.t.c. are not available for communication. Hence, this study will fill this gap by ensuring that provision is made for communication in about **1** to more than **50** languages and particularly incorporating African languages such as Kinyarwanda, Swahili, Hausa etc. A key benefit of this development is that it will promote the popularity of native African languages and their knowledge of local cultures and secondly, the communication will be more convenient for a wider client population, particularly the locals in African regions.

In the discussions presented by Bharti et al., (2020), it was deduced that a key shortcoming associated with the virtual medical consultant is the lack of empathy and visceral connection between patients and the virtual medical consultant. Although patients received prompt responses from the system during the consultation, the feedback from the system indicated a lack of satisfaction characterized by the absence of empathy. From the study, it was inferred that patients felt that there was no display of kindness or love and that the interaction felt like a business discussion. This is an essential limitation that characterizes the contemporary conversational AI systems, and attempts will be made in this research to rectify this flaw.

The research by Razumovskaia et al., (2021) pointed out that the major challenge confronting the advancement of chatbot systems is the unavailability of data in not so popular languages particularly, African languages. This constraint makes it difficult to deploy and train machine learning models that can understand and interpret these local languages. In considering alternative solutions to this problem, the use of AI language translators to translate datasets in popular languages to local dialects is a viable solution that was not captured in the research. Hence, in this research, the use of language detectors and translators will be instrumental in providing data in unpopular/local languages in the form of detecting and translating users' queries and responses.

## **2.5 Research Contribution**

This study proposes and implemented an NLP-based multilingual conversational AI bot that offers educational information provisioning services and guidance to students and the general public in 1 to over 50 languages; the bot is equipped with empathy and intelligence capabilities to converse with the general public in multiple languages inclusively African languages such as Kinyarwanda, Swahili, Hausa, Igbo etc. This software application aggregates the services of a consultant by offering recommendations, Q & A to FAQs, and the most potent educational information. Thus, the study proposes a novel AI solution that serves as a client's virtual consultant/assistant. The study program is carefully created and the machine learning model thoroughly trained to communicate with clients intelligently as if it was a real human server.

## CHAPTER THREE: RESEARCH DESIGN AND METHODOLOGY

### 3.1 Introduction

This chapter seeks to provide the logical procedures and methods employed in the study in order to accomplish the study's aims and objectives. It covers the methodology, research design, the research procedures, data collection method, data analysis method, validity and reliability, data preparation method, Machine Learning models, and the technological tools to be utilized for this work.

A mix of ML, NLP, NLU, and/or transfer learning techniques can be used to develop chatbots in general (Alruqi & Alzahrani, 2023). The multi-lingual AI chatbot in this study integrates ML, NLP, NLU and in addition language detection and translation mechanisms. Thus, the operation of the chatbot for this study is divided into two distinct parts: The user query bot, and the language detection and queries and predictions translation mechanism. The software application was built using Python 3.9 and Flask Framework. The Gated Recurrent Units (GRUs) ML model for the study was built on the Pytorch deep learning framework. The language detector and translator used are langDetect and Google translator APIs. The implementation was done using Visual Studio Code Editor.

The two methodologies (the Agile and Cross Industry Standard Process for Data Mining (CRISP-DM) methodologies) are deployed in this work to serve distinct purposes. The CRISP-DM is deployed to foster the research and evaluation of machine learning models to be subsequently utilized in the software development phase. It serves as a framework of procedures or operational structures employed as guidance for developing a comprehensive Artificial Intelligence/Machine Learning project (Pradana and Rochmanto,2022). This choice is primarily because the CRISP-DM is a foundational process model that forms the basis for a data science process, thereby extensively informing the choice of most suitable models for different context with critical justification. The CRISP-DM captures the research design (i.e. Action Research), it is further used to ensure and evaluate the reliability and validity of the data collection instruments, the sample size and sampling techniques, the data collection procedure, and method of data analysis. The CRISP-DM would essentially foster the attainment of the predefined objectives and ultimately provide answers to the research questions.

Because this project comprises both theoretical (research) and a practical component (software development), the need for a methodology most appropriate for the practical component led to adopting the Agile (Kanban) methodology for the implementation phase. The Agile (Kanban) methodology is deployed for the implementation/development of the software as it encapsulates standard industry

practice and methods, it is used to plan and direct the development process, bringing it to fruition promptly and efficiently. The choice of the Agile methodology for the implementation component is essentially because it ensures pragmatism and efficiency by promoting flexibility, adaptability to changing requirements, and consultation/collaboration which target producing standard software in a short time span.

Lastly, the combination of the CRISP-DM and the Agile methodologies will yield the best results because they both place strong emphasis on reliability and adaptability. The CRISP-DM outlines the standard AI project process flow which encapsulates the research design– data collection methods, Data analysis methods, whereas the agile methodology defines the structure of the implementation phase, the individual tasks, resources, deadline, as well as the procedures and tools required for execution.

### **3.2.1 The Agile Methodology**

The Agile methodology is an interdisciplinary approach to project management that provides an iterative technique to manage projects and software development using test-driven design and ongoing input to address difficulties. It is a set of best practices that constantly include planning, collaboration, and reviewing of progress on the go. According to the study of Michael et al. (2021), the agile methodology emphasizes the iterative development of functional components in little increments during the project implementation. Hence a key advantage of using the agile methodology for this study is the divide-and-conquer-approach as described which fosters efficiency, effectiveness, prompt delivery, and flexibility which essentially promotes the creation of comprehensive, sophisticated, and trustworthy transdisciplinary frameworks/models. Additionally, the agile methodology offers companies and developers adequate time to produce and sell their goods successfully whilst concurrently handling users' changing wants on the go. That is, the agile methodology is particularly suitable for this project because it allows the execution of multiple stages and tasks simultaneously rather than sequential execution approach as observed in the waterfall methodology, hence, this will hasten the project development while maintaining the required level of quality and procedures.

A study by Khaza and Syam (2020), affirm that the ability of a project to handle recurrent changes is essential to its success, and the agile methodology is defined by its dynamic in task management and its flexibility to occasional modifications that advances the project. These features of the Agile methodology are relevant to this project because streamlining the development process into smaller segments will simplify the task of overseeing the project as a whole and track improvements. Hence, this is the justification of the agile methodology as the choice methodology for this project as it enables

easy integration of different components throughout the software development process and also enables quick responsiveness to changes in client needs.

The agile methodology improves structure and production value in addition to its scalability and reliability. Hence, the fact that this project is time-constrained and an AI project, these necessitates a rapid development workflow and the process for implementing a Data Science based project must be followed critically, are additional important aspects that make the agile methodology a preferable choice. Figure 3.1 below depicts the Agile Methodology.

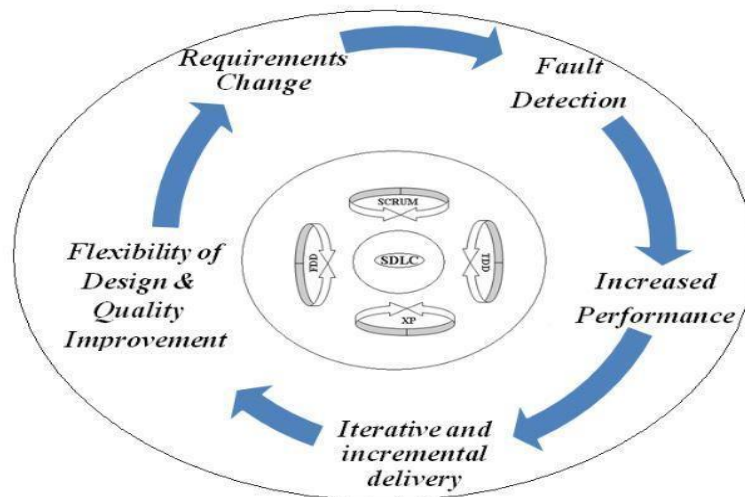


Figure 3.1: Agile Process flow (Gaurav & Pradeep, 2012).

According to the study of Gaurav and Pradeep (2012), the significance of each of the items in the aforementioned Agile process flow are described as follows:

- i. **Performance Improvement:** Participants in a project can share useful knowledge and iteratively improve things by holding regular meetings in Agile project process. This can result to increase in team productivity, the overall return on investment can be higher than the sum of the individual results. Therefore, through multiple meetings between the researcher and the supervisor, where the supervisor's suggestions and advice were accepted and put into practice, performance improvement on this project was achieved.
- ii. **Handling Change in Requirements:** Since clients are active participants in the development process of every standard project, i.e., clients could efficiently manage project operations through on-site engagements and input from meetings. In the agile methodology, the planning process of a project is significantly improved because the current requirements accurately reflect the current and latest demands of end-users. Following the Object-oriented paradigm's traditional software development approaches made it easier to handle changes in customer needs. Additionally,

complete documentation of all user requests and actions conducted during the project was completed and was periodically referred to during its entire life cycle of the project.

- iii. ***Delivery that is iterative and incremental:*** To optimally minimize risks and obtain early client feedback, projects are divided into smaller units with discrete functional deliverables. This component was implemented in the work by delivering and testing each completed unit at each stage of the project to allow for evaluation and improvement before it was finally integrated into the whole chatbot system.
- iv. ***Improvement in Quality:*** Project refactoring and test-driven development techniques are used to improve project quality. Refactoring, makes the source code more reusable and enhances the quality of a software development. Thus, every aspect of the software component is improved, including the design, architecture, and output quality of each system unit of the chatbot.
- v. ***Flexible Design:*** In project implementation, flexibility refers to the ability to quickly change or shift course swiftly during implementation. To achieve this, the software design needs to be flexible enough to handle requirements modifications since this is a crucial aspect of the Agile methodology.
- vi. ***Fault Detection:*** In the agile methodology, testing is done in every iteration of the project implementation to detect defects early and be able to rectified them before they may become more serious, as opposed to a plan-driven process approach. Additionally, continuous testing makes it possible to test input continuously, which helps to enhance code in later versions.

The components in figure 3.1 above depicts the major phases that were adopted to guide the direction and overall development process of the Agile methodology. The project plan detailing the timeline of the constituent activities required was designed using the Trello board, as depicted in figure 3.2 below.

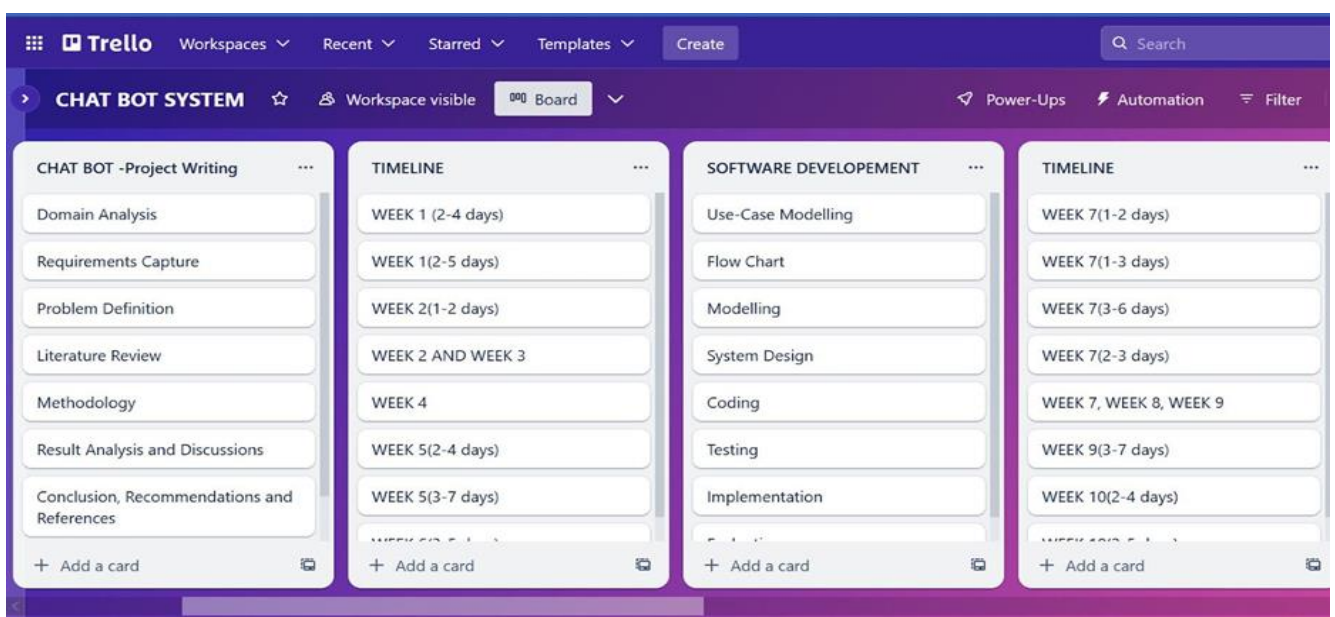


Figure 3.2: Project Plan/Management Using Trello Board

### 3.2.2 Research Design– Action Research

The research design aligned with the CRISP-DM methodology is Action Research. The Action research, as emphasized by Davison, Martinsons, and Kock (2004), enables the investigation, uncovering of issues and the implementation of practical solutions in the real world. The goal of action research is to simultaneously research, examine, and solve a real-world problem. Hence, action research combines conducting research accompanied by actual action, as the name suggests. The Action research was chosen because it involves addressing a real-world problem, documenting the solution process, and developing an iterative action plan to achieve best practices which is most applicable for this work. Furthermore, the Action Research is most appropriate for this project as it ensures that by the end of the research endeavour, a practical and workable solution is developed for a real-world problem and deployed. The Action Research design is suitable for this project owing to the following additional reasons:

- i. **Pragmatism:** Action research is mainly motivated by resolving practical concerns and real-world problems. Hence, it ensures that researchers concentrate on pertinent and meaningful information or discoveries that can be applied to the project to make real-life advancements or produce practical solutions.
- ii. **Iterative:** An iterative cycle of planning, doing, observing, reflecting, and further planning characterizes action research. This cyclical process enables researchers to constantly improve and adapt their procedures, designs, and solutions in response to input and novel findings, resulting in higher-quality results.

Key components of the Action research are outlined in the CRISP-DM project process below:

Figure 3.3 below presents the CRISP-DM AI/ML project process.

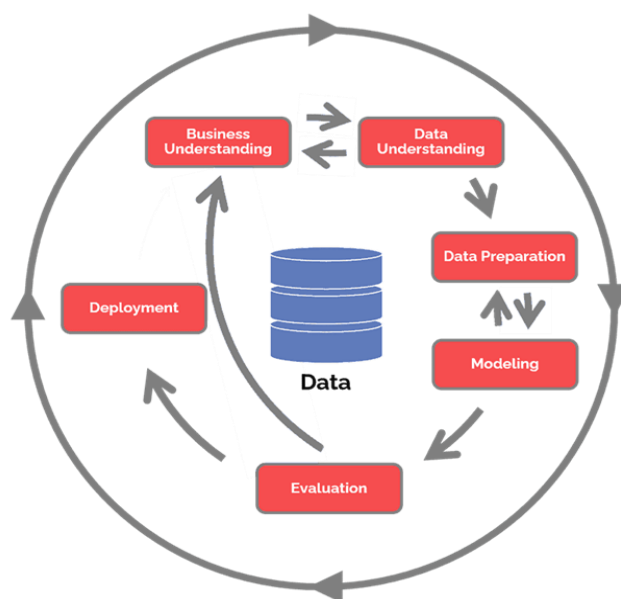


Figure 3.3: CRISP-DM Diagram (Data Science Process Alliance, 2023).

The description of each phase of CRISP-DM as it pertains to the implementation of the Multi-Lingual Conversational AI Chatbot of this study is detailed below:

### **i. Business Understanding**

In this study, understanding the particular needs of ACEDS, University of Rwanda, as well as the difficulties it has with regards to information acquisition by the general public or visitors within the educational sector, served as the driving force that birthed this study. The inquiry was motivated by the decreasing availability of first-hand information or real-time consultation channel for international or prospective students and the general public who might not have direct access to ACEDS and its management or feel uncomfortable communicating exclusively in English or Kinyarwanda. The business understanding phase involved a number of activities, starting with an evaluation of ACEDS overall information acquisition by the general public, followed by an analysis of this area to pinpoint any issues that were identified. This is carefully followed by plans and strategies to develop and overcome these problems using data mining techniques.

### **ii. Data Understanding**

Data understanding is essential for creating an efficient Chatbot system utilizing language models and machine learning technologies. Therefore, the quality of the training datasets for a chatbot system has a significant impact on the chatbot's success because it ensures that only the most relevant data types are used to train the model. Hence, suitable type of data relevant for training the Chatbot system for this study can be categorized into the following:

- a) **Knowledge data** – knowledge data are key domain specific information that educate the general public on a subject matter or about organizational or industry services, it provides guidelines or instructions on solving a problem or to certain information. Knowledge data are supplied to queries that typically comprises Frequently Asked Questions (FAQs) submitted by clients or employees in addressing issues, e.g. programs information's, scholarships, internships, study fees, registration processes, admission, accommodation, events, training facilities, campus club, sport and so on as it pertains to this study.
- b) **Conversational data** – Conversational data are basically data that could be obtained from interactions between two or more parties, and can be obtained from a variety of sources, including real-life chat logs, phone calls conversations, customer service contacts, social media conversations, and even manually generated conversations, or publicly available datasets

specifically designed for the purpose of chatbot training. This data is used to train and improve the chatbot's NLP capabilities, and importantly, allow the chatbot to understand user queries and predict appropriate responses.

- c) ***Sentiment data*** – Sentiment data refers to information that captures the emotional tone expressed by users during interactions with artificial agents, e.g. with the chatbot system. It involves analyzing the sentiment of user inputs, to understand the user's mood, i.e. if they are positive, negative, or neutral at a moment. This data can help in assessing user satisfaction and understanding the effectiveness of the chatbot's responses. In this research, open source sentiment data from Kaggle was used together with other data to train a machine learning model for the purpose of improving the chatbot's performance, and identify areas for enhancement, and ensure better users and system engagement and satisfaction.
- d) ***Contextual data*** – Contextual chatbot data refers to the relevant information and context specific information to a conversation. It encompasses the user's past queries, system responses, preferences, and session history of users and systems conversation. This data is very important because it helps the chatbot maintain continuity, personalize responses, and provide accurate information based on users' sessions history. It ensures a seamless conversation and improves the user experience while engaging the chatbot system in a conversation. Therefore, contextual data provides a deeper understanding and helps analyze client's behavioral patterns for enhanced experiences and engagement. **Note:** This context is not captured in the training data of this research due to the following reasons: there were not enough volunteers willing to produce real-time data and the time allotted for the completion of the project was short to explore measures appropriate for implementing this context.
- e) ***Annotated data*** – Annotated chatbot data can be refers to as the type of dataset that has been manually labeled or automatically using machine learning algorithms or annotated with additional information to provide specific insights or instructions to the chatbot. This data typically includes annotations such as intent labels, entity tags, sentiment labels, or dialogue acts for consultation purpose in a certain domain. Annotated data can help train and improve the chatbot's understanding and intelligence in accurate response provisioning. It serves as a valuable resource for machine learning algorithms to learn patterns and make accurate predictions in future interactions between users and the system. It plays a crucial role in developing, fine-tuning and deploying chatbot models for better performance and more effective communication with users. These are applicable in the context of this research by making sure all the collected data for the model training were annotated into different class labels before fitting the deep learning algorithm.

### iii. Data Collection Methods

Data for this study were FAQs and observations collected from the AC-EDS Staff and the ACE-DS website(<https://aceds.ur.ac.rw/>) e.g. fees, programs, facilities, certifications programs, and staff profiles with links to useful information. Additional chatbot data(Conversation and sentiment data) were obtained from the open-source chatbot Kaggle dataset (<https://www.kaggle.com/datasets/niraliivaghani/chatbot-dataset>) to improve the chatbots' functionality and increase users' satisfaction. The dataset collected and annotated contains 88 tags(topics) and 1356 patterns. The presentation in Figure 3.4 below is a snapshot of the annotated data for the chatbot model training showing establishment tag.

```
1  {
2    "intents": [
3      {
4        "tag": "establishment",
5        "patterns": [
6          "what year was ACEDS established?",
7          "How old is ACEDS?",
8          "how many years is ACEDS",
9          "year of establishment of ACEDS",
10         "The year ACEDS was founded",
11         "establishment",
12         "ACEDS was founded in the year",
13         "founded?"
14       ],
15       "responses": [
16         "ACEDS was Founded in september 2017",
17         "ACEDS was established in september 2017"
18     ],
19     "context_set": ""
20   },
21 }
```

Figure 3.4: Chatbot Dataset Snapshot

### i. Data Preparation

The Multilingual Conversational AI Chatbot was developed and made intelligent with an annotated data used in training a deep learning algorithm (GRUs Recurrent Neural Network). The phases in the data preparation are captured in the presentation in figure 3.5 below. These stages include determining the chatbot's target, purpose, and capabilities; collecting relevant data; categorizing the data; annotating the data; balancing the data; regularly updating the dataset; and testing the dataset by training a machine learning algorithm and evaluating the performance. These preparation stages are an iterative process and requires on the go attention to maintain the chatbot model's relevance and effectiveness.

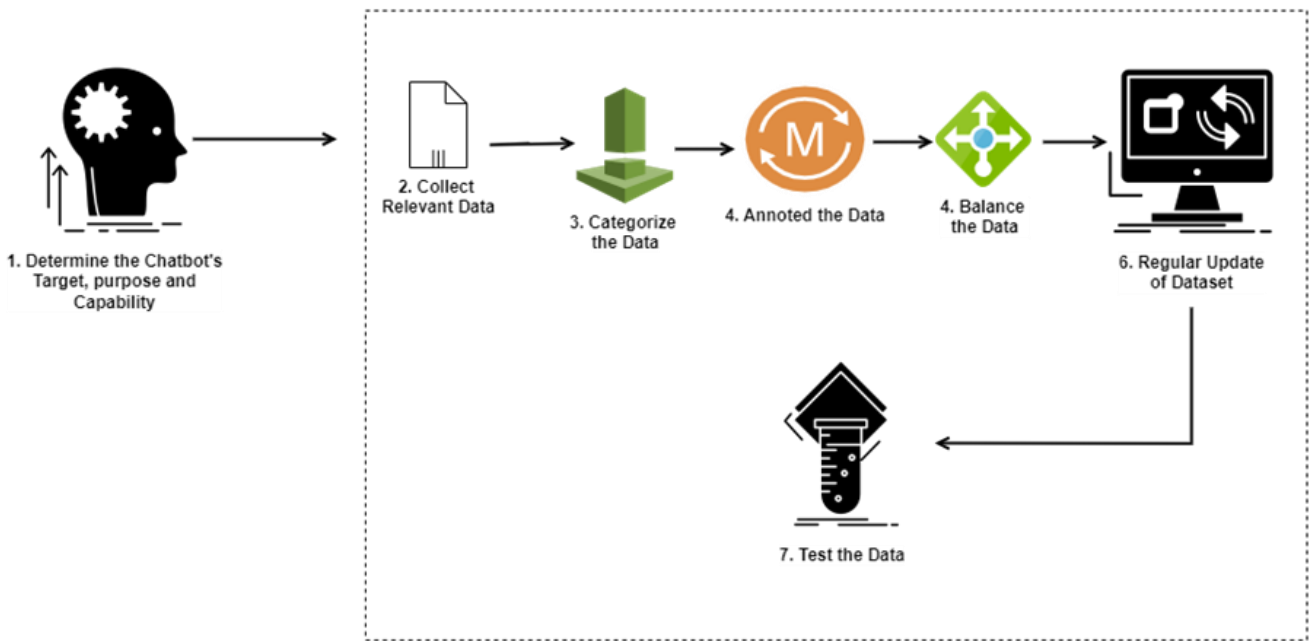


Figure 3.5: Data Preparation Diagram

Figure 3.5 above depicts the different stages of data preparation, which are indispensable for ensuring a well-formatted and enhanced chatbot training data. The descriptions of the data preparation phases are briefly explained as follows:

**a) Chatbot System Target and Capability**

To ensure a high quality chatbot dataset, the following factors must be considered:

- i. **Purpose:** This involves collecting relevant data and design conversation flow based on the chatbot's specific tasks. In the context of this research, the target is for educational information provisioning for the purpose of studying at ACEDS. The capabilities include FAQ and Answers provisioning, amongst others.
- ii. **Medium:** This involves choosing the appropriate training data based on the chatbot's medium, whether the data is voice or text based. In the context of this research, its text based.
- iii. **Languages:** This involves incorporating multilingual data into the dataset if the chatbot needs to support multiple languages for its consultation services. In the context of this study, language recognition and translation APIs were utilized to identify and translate users queries and responses into required languages, facilitating the language translation processes.

## **b) Acquire relevant data**

The next step involved is to gather relevant data that is pertinent to the chatbot's project domain. Data collection is a crucial and indispensable phase in dataset preparation as it provides the raw materials for training and fitting the AI model with the intelligence it needs. Therefore, the chatbot training dataset should consist of a combination of primary and secondary data.

## **c) Categorize the data**

Once the data is collected, the next is for the data to be organized into categories based on topics and intents or tags comprises of classes, and patterns. This can be achieved manually or with the assistance of NLP algorithms. Categorizing the data allows for a well-structured/formatted data, arranged in a way capable of enabling the chatbot system to learn and identify the underlying patterns, specific topics and intents in the dataset. For instance, in this study data were categorized into topics such as About ACEDS, hostel, English language requirements, application procedures, master's programs, and so on.

## **d) Annotate the Data**

The critical step that comes after categorization is done, is data annotation, often known as dataset labeling. This is done in order for conversational AI models, like chatbots and virtual assistants, to comprehend the intent and significance of client messages. Annotating or labeling of the chatbot training data are essential in the context of this research. The data for this system were manually annotated.

## **e) Balance the data**

Achieving a balanced and comprehensive dataset is crucial to preventing models' bias in the chatbot's performance because unbalanced data leads to model overfitting. The dataset should encompass all relevant topics and intents, and should be well balanced, enabling the chatbot model to effectively respond to a wide range of user requests.

## **f) Regularly update of dataset**

For any sophisticated Chatbot solution, the Chatbots data requires frequent updates to maintain optimal performance and up to date information provisioning. That is, new data, including: client's interactions, feedback, and changing trends, should be collected, analyzed and incorporated into the model data. This ensures the chatbot stays up to date, relevant, adapts to evolving user expectations, and learns new response variations for better service provisioning.

Thus, regularly updating of chatbot dataset helps chatbots deliver improved, up-to-date experiences for its users.

#### **g) Test Dataset for Accuracy**

This phase involves training the chatbot with a smaller subset of the whole dataset and evaluating its performance on unseen data. That is, some portion of the data the model has not be trained with, and by doing so, any gaps or shortcomings in the dataset can be identified, leading to a more effective chatbot system.

### **3.3 Modeling**

Predictive models are constructed using real-world experiences and information for training an AI model which critically learn the underlying patterns in the information provided to the model to accurately make predictions based on knowledge scope of the fitted data. Predictive models are utilized as an artificial agent to simulate real world events. According to Dreiseitl et al. (2001), models can be knowledge-based expert systems with predetermined rules or training data for statistical and machine learning models. The Multilingual Conversation AI Chatbot Model configuration makes used of a Gated Recurrent Units (GRUs), an advance kind of RNN architecture that have LSTMs features and in addition, have information flow controlling gating techniques that make them computationally more efficient than LSTMs and still captures the long-term dependencies in LSTM architecture. The following are taken into consideration in building the GRUs Recurrent Neural Network for this study:

- i. ***Problem Definition:*** This necessitates a precise problem definition of the tasks that the Multilingual Conversational AI system will perform, which is a classification problem because the model will be able to recognize the category or intent of the user's inputs. In order to further improve its ability to deliver pertinent responses, the chatbot can categorize user inputs into predetermined intents or tags or classes.
- ii. ***Data Preprocessing:*** Before fitting a model with training data, the dataset must be cleaned, normalized to the same casing, tokenized, lemmatized, and added to word list known as bag of word in NLP, and all unique words sorted and transformed to a suitable format (NumPy array) for training the GRU RNN Machine learning algorithm. Various NLP python libraries such as nltk and spaCy toolkits where used for the data preprocessing.
- iii. ***Network Architecture:*** Network Architecture was determined to allow the model have a deep understanding of the fitted training data. In the context of this research, the model architecture

includes 3 numbers of layers which include the input layer, hidden layer and the output layer; including are the numbers of neurons in each layer, and the connections between the layers.

- iv. **Activation Function:** The activation function selected for the chatbot model configuration is rectified linear unit (ReLU). ReLU is a deep learning activation function that imparts non-linearity to a deep learning model and addresses the problem of vanishing gradients when the algorithm is learning the patterns in the data. It focuses on capturing the positive aspect of its input.
- v. **Training Algorithm:** The training algorithm chosen to update the network's weights and biases via an iterative process is backpropagation.
- vi. **Training Process:** The GRUs RNN model was trained using the prepared intents by using the backward propagation training algorithm to adjust the weights and biases iteratively until the network reached a satisfactory level of accuracy and loss value.
- vii. **Optimization and Fine-Tuning:** Hyperparameter adjustment was done and the Adam optimizer is chosen as a regularization technique to improve the performance of the GRUs RNN deep learning model.

After the configuration of the GRU RNN, the modeling process is as presented in the figure 3.6 below.

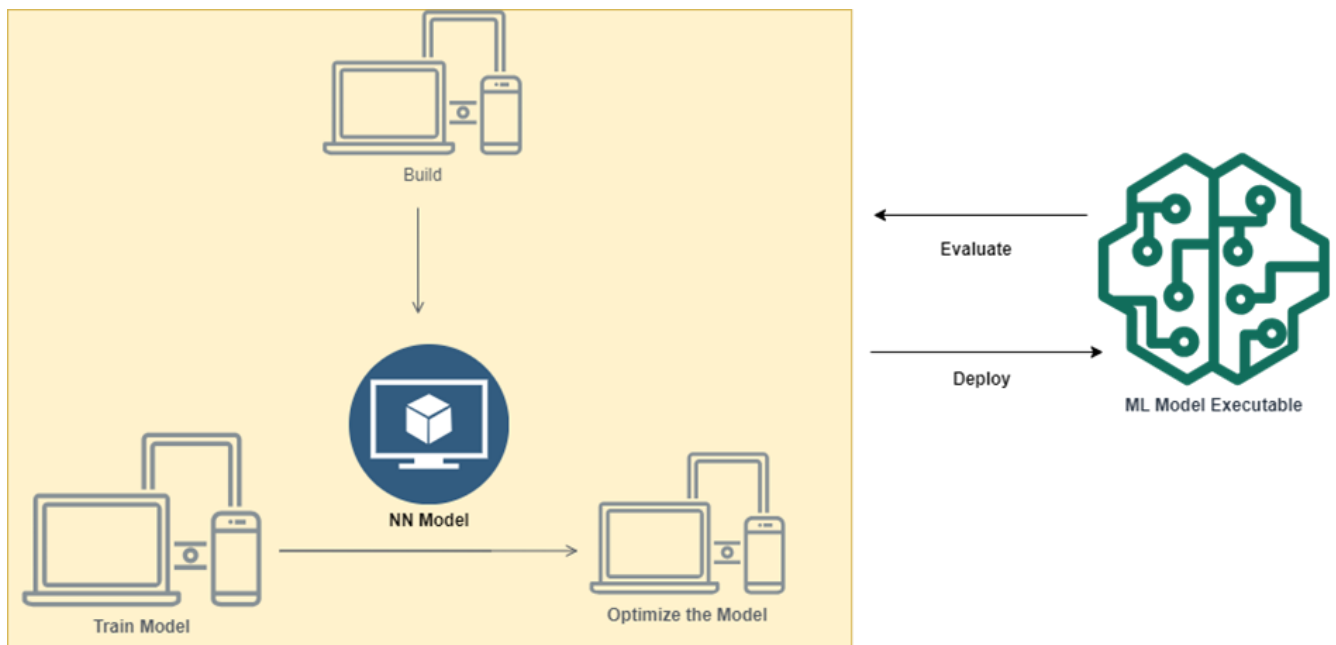
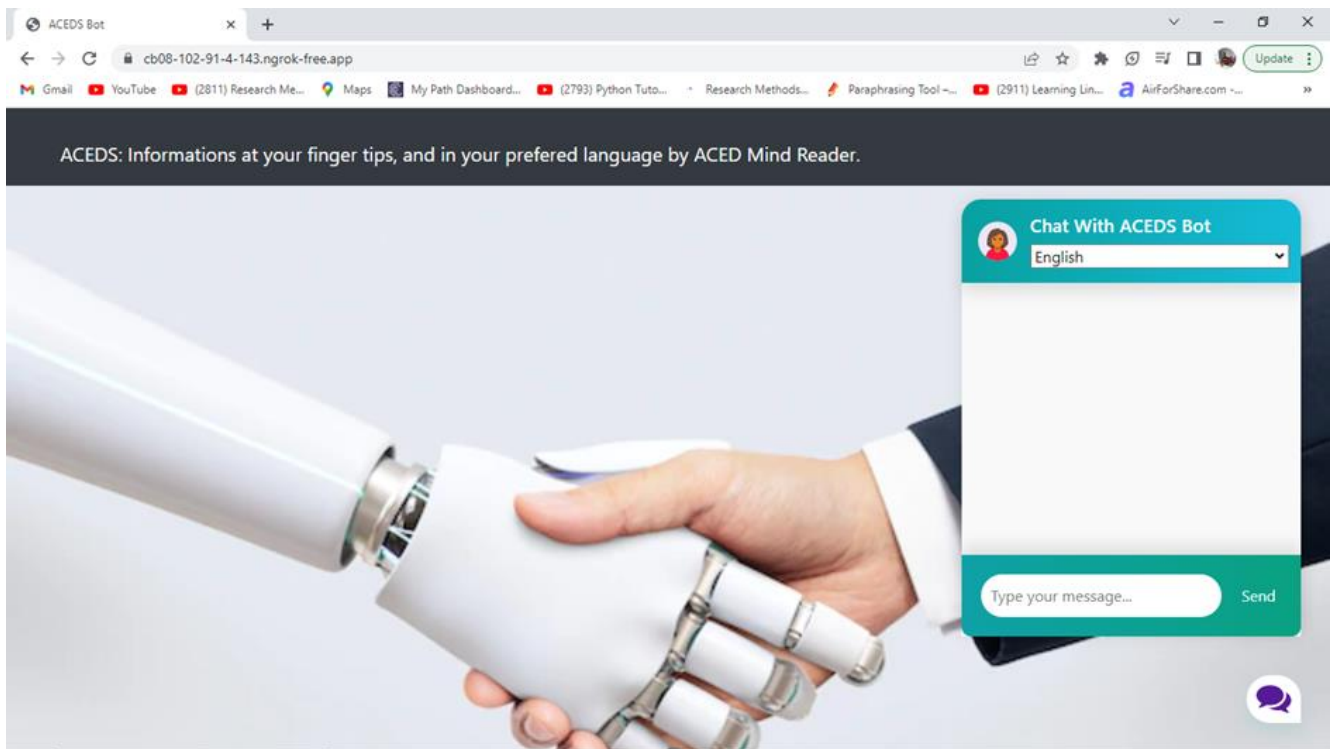


Figure 3.6: Modeling Presentation Diagram

Figure 3.6 above presents a high-level representation of how the GRU Neural Network Model. The model was configured, trained, and fine-tuned on a local personal computer (PC). During training, the Chatbot model achieved a 98% accuracy with a loss of 0.000 after 1500 epochs. The trained model is saved after the 1500 epochs, and is integrated into other component of the software abstracted by the

user interface. Then hosted on the local host server. It was tested through model validation and real-time testing, resulting in excellent prediction results for users queries with reference to it trained knowledge scope. To make it accessible online, NGROK, a utility for creating secure public tunnels to locally hosted applications, was used as a reverse proxy to make the application accessible via the internet by the public. This allowed the Multi-lingual AI Chatbot system to be accessed and tested by the public via web browsers like Edge, Internet Explorer, or Google Chrome. The Multi-lingual Chatbot's Graphical User Interface (GUI) is presented in the snapshot in *Figure 3.7 below*.



*Figure 3.7: The Multilingual Chatbot GUI Presentation*

### **3.3.1 Testing Techniques Deployed to Determine the Validity and Reliability**

Software testing is a crucial stage in the creation of software solutions that makes sure the features of the program comply with the specified goals of software implementation (Vahid, Michael, and Feyza, 2019). It entails evaluating and confirming that the software's effectiveness in achieving its objectives is successfully accomplished. According to Mahboubeh et al. (2020), testability describes how effectively a software system or unit enables testing. Software development requires quality assurance to guarantee dependability and specification compliance of the project. The study of Mahboubeh et al. (2020), verification and validation procedures seek to satisfy users' needs and develop applications that adhere to specifications. According to the study of Vahid et al. (2019), software testing is a key and extensively used technique for verifying the quality of applications effectiveness. In this study, rigorous software

testing was done to verify the system in order to accomplish a standard software expected for the purpose of the study.

Different testing techniques are employed to determine the validity and reliability of the solution in this study, for example, unit testing, user testing, real-time testing was carried out during the implementation of the chatbot system. Unit testing, according to Mahboubeh et al. (2020), evaluates numerous components of the software system independently in order to ascertain its workability. In this study, unit testing of the software is performed to ensure that the software components are fully functional at an independent level so as to ensure manageable transition and integration of the units into a comprehensive whole system. Integration testing was carried out to ensure the correct functionality of the whole system is achieved and working well, high-level testing was carried out to evaluate the system's behavior and overall real-time performance when deployed to real-life scenarios, and users testing was undertaken and users' feedbacks obtained to ascertain the user's friendliness and satisfaction while conversing with the chatbot. These testing steps were carried out concurrently with the development phase in the agile methodology and CRISP-DM process used for this project, resulting in a resilient system with fully functional and effective modules.

### **3.4 Evaluation Metrics**

The evaluation process for the multilingual conversational AI chatbot system encompasses training and validation evaluation, real time testing evaluation, users testing evaluation. The evaluation metrics utilized the Scikit learn(sklearn.metrics) package consisting of the classification report and confusion matrix, and confidence level which automatically evaluates the models performance and predications of users queries. The chatbot solution was also evaluated based on its ability detect languages, translate and process unstructured user input messages and responses to users target language. This evaluation involved tasks such as extracting entities and determining the intent of the user's message based on predefined categories in the knowledge base of the chatbot model.

#### **3.4.1 Evaluation of the Dialogue Model**

The dialogue model was evaluated based on its ability to accurately generate appropriate responses to user inputs messages or queries during a conversation on the user interface (UI) of the chatbot. The evaluation considered the accuracy of the responses in relation to the understanding of the message and the intent classification achieved by the NLU model. Since the chatbot solution is a classification problem, the evaluation process involved creating a Classification Report detailing the Accuracy, Precision, F1-Score, Recall, Support, weighted average of precision, recall and F1-score, and the

Confusion Matrix table, which has the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values of the models' metrics report on the chatbot performance on the categories of tags in the training data. These components were used to calculate the following evaluation metrics:

- i. **Accuracy:** The accuracy parameter of the model metric report measures the overall accuracy rate of the model's predictions or classification performance rate. Accuracy is a metric that measures a model's ability to predict class(tag) labels correctly. That is, it represents the average accuracy of the classifier's predictions of the model class. Mathematically, **Accuracy = (TP + TN) / n**, where **n** is the total number of observations (Kohavi & Provost, 1998).
- ii. **Precision:** Precision is an evaluation metric that indicates the accuracy rate of classification results in the systems' predictions of a tags. It evaluates the model's accuracy and quality. Mathematically, **Precision = TP / (TP + FP)** (Kohavi & Provost, 1998).
- iii. **Recall: Recall also known as True Positive Rate (TPR) or Sensitivity:** This indicates the success rate of the system in recognizing a particular category(tag). That is, it is a measure of the classifier's ability to correctly predict positive instances. It quantifies the scale or comprehensiveness of the predictions of the model. **Recall or TPR or Sensitivity = TP / (FN + TP)** (Kohavi & Provost, 1998).
- iv. **F1-Score:** The F1-score provides a description of the relationship between precision and recall, often referred to as the harmonic mean. The F1-score is a measure of a model's accuracy that considers both precision and recall of the model. It is referring to as the harmonic mean of precision and recall, providing a balanced evaluation of the model's performance. **F1-Score = 2 X (Precision X Recall) / (Precision + Recall)** (Kohavi & Provost, 1998).
- v. **Support:** Support in the metric report denotes the count of instances or the samples belonging to each class in the model. In a nutshell, it quantifies the number of true occurrences associated with each class within the dataset. The support value offers valuable insights into the distribution and representation of individual classes, enabling a deeper understanding of the reliability and significance of the evaluation metrics for each class. **Support = TP + FN** (Kohavi & Provost, 1998).

**Confusion Matrix:** A confusion matrix, as defined in a study by Kohavi and Provost (1998), is a tabular representation that assesses the performance of a machine learning classifier algorithm. It presents information about the classifier's projected and actual classifications. The confusion matrix is an evaluation metrics widely used to evaluate the performance of a classifier model. Figure 3.7 below detailed the parameters of a confusion matrix.

$$\text{Confusion matrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

Where,

“a” is the  $N_0$  of accurate predictions that a given instance is “-”,

“b” is the  $N_0$  of incorrect predictions that a given instance is “+”,

“c” is the  $N_0$  of incorrect of predictions that a given instance “-”,

“d” is the  $N_0$  of accurate predictions that a given instances “+”.

*Figure 3.8: Confusion Matrix (Kohavi and Provost, 1998)*

**Classification Report:** Classification Report was particularly chosen for the model evaluation of the model performance, and also used for documentation of this study evaluation results because, a classification report provides a more detailed information about model performance. It provides metrics such as the precision, recall, F1-score, support, and their weighted average for each class in the classification task. It presents these metrics for both individual classes and provides an overall average across all classes. The classification report is a better evaluation metric especially when dealing with multiclass classification problems as in the case of this study with up to 88 classes, where the confusion matrix may become more complex to interpret due to the numerous classes. It provides a concise summary of the model's performance and can help in comparing different models or variations of the same model (Kohavi & Provost, 1998). While a confusion matrix provides a visual matrix representation of the model's predictions, a classification report offers a tabular information with more comprehensive and informative evaluation of the model's performance, making it a suitable choice for reporting and documentation purposes.

### 3.5 Deployment

The implementation of the Multi-lingual Conversational AI Chatbot utilizing GRU version of RNN incorporates a user-friendly Graphical User Interface (GUI) to enhance user satisfaction in accessing the chatbot solution. The diagram in figure 3.9 below illustrates the implementation process of the chatbot service system for end users.

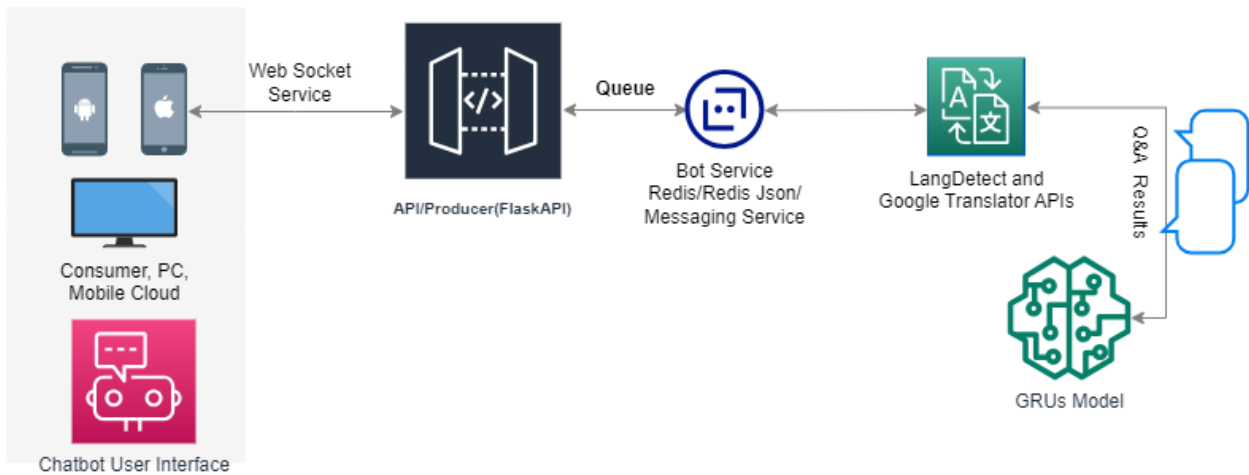


Figure 3.9: The Deployment Architecture of the Multi-lingual AI Chatbot Solution

In the presentation in figure 3.9 above, Language Understanding and Translation APIs were utilized to detect users query language, users target language and translate the information processed into required languages, thereby facilitating the translation process.

### 3.5 Limitations: Shortcomings in the Methodologies.

Some of the shortcomings of the Cross Industry Standard Process for Data Mining methodology observed during the course of this research encompass the following:

- i. Although the CRISP-DM places greater emphasis on the creation of data mining techniques, it fails to provide detailed instructions on how to set up and sustain them in a real-world setting. This is important for projects and businesses that are long-term oriented and require regular upgrade and maintenance.
- ii. On activities involving data prep and gathering, CRISP-DM offers only rudimentary guidance. It makes the presumption that the data is immediately accessible and in an appropriate form, which may not always be possible in actual-life circumstances.

*Some of the limitations of the Agile methodology include the following:*

- i. The excessive flexibility involved often fosters unpredictability because there is no rigid structure and definite order of task execution.
- ii. It creates greater workload on the developers because at any point in time, new changes or updates may be requested by clients and must be incorporated no-matter the stage which the project is.
- iii. At any point in time, it is difficult to evaluate the whole project because of the flexibility involved where team-members sporadically switch between tasks.

### **3.6 Development and Learning Tools**

**3.6.1 Implementation, Tools and Materials:** The system was developed, deployed and locally tested on Windows 11 PC, 8GB RAM, using the appropriate implementation environments and technological tools, such as Visual Studio Code and Google Collab. The Chatbot was built with Python programming language and Flask web framework RestFul APIs. A Bootstrap framework made up of HTML, CSC, and Javascript was used to design and develop a resopnsive user interface (UI). Natural Language Text Preprocessing Techniques such as data normalization, tokenization, lemmatization and bag of words was used to prepare the data (Intent) for training using Feed-Forward Neural built on the PyTorch Deep Learning Framework. Google Translator and LangDetect APIs were used for language detection and queries translations. GRUs RNN deep learning algorithm was trained using the study data.

**3.6.2 Technological tools: Backend:** Python3, Deep learning Framework (Pytorch), RestAPI(Flask), Language Detection and Translation APIs, Natural Language tool kits(NLT) libraries, and Neural Network.

**Frontend:** Bootstrap (HTML, CSS, JavaScripts.)

**Development tools (editors):** Visual Studio Code Editor and Google Colab

## CHAPTER FOUR: SOFTWARE TESTING, RESULTS, EVALUATION AND DISCUSSION

### 4.1 Introduction

This phase of the study encompasses the results of the methodology; software development results, real time testing analysis results, discussion, user's evaluation feedback, and evaluation of the study solution as a whole. The performance of the chatbot is summarized and presented in a tabular and visual format, showcasing the system testing, training and validation evaluation performance, and pie charts for the user's evaluation feedbacks. Descriptive analytics play a crucial role in order to effectively communicate information about the behavior and performance of various components of the chatbot system and its evaluation.

### 4.2 Test Analysis

#### 4.2.1 Unit Testing

An illustrative instance of unit testing is observed in the model training process conducted using the deep learning neural network algorithm to get an effective and functional model. The model training is assessed and validated using various parameter values to finetune the model prior to its integration into the main chatbot system. The code listing below depicts a unit testing implementation of the GRUs RNN model training.

```
# Hyper-parameters
num_epochs = 1500
batch_size = 8
learning_rate = 0.00134
input_size = len(X_train[0])
hidden_size = 8
output_size = len(tags)
print(input_size, output_size)
class ChatDataset(Dataset):
    def __init__(self):
        self.n_samples = len(X_train)
        self.x_data = torch.tensor(X_train) # Convert to torch tensor
        self.y_data = torch.tensor(y_train, dtype=torch.long)# Convert to torch tensor

    # support indexing such that dataset[i] can be used to get i-th sample
    def __getitem__(self, index):
        return self.x_data[index], self.y_data[index]

    # we can call len(dataset) to return the size
    def __len__(self):
        return self.n_samples

dataset = ChatDataset()
train_loader = DataLoader(dataset=dataset,
                           batch_size=batch_size,
```

```

        shuffle=True,
        num_workers=0)

device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
model = NeuralNet(input_size, hidden_size, output_size).to(device)

# Loss and optimizer
criterion = nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.parameters(), lr=learning_rate)

# Train the model
correct_predictions = 0
total_samples = len(dataset)
loss_values = []
all_predicted_labels = []
all_true_labels = []

for epoch in range(num_epochs):

    for (words, labels) in train_loader:
        words = words.to(device)
        labels = labels.to(dtype=torch.long).to(device)

        # Forward pass
        outputs = model(words)
        # labels = torch.max(labels, 1)[1]
        loss = criterion(outputs, labels)

        # Backward propagation and optimization
        optimizer.zero_grad()
        loss.backward()
        optimizer.step()

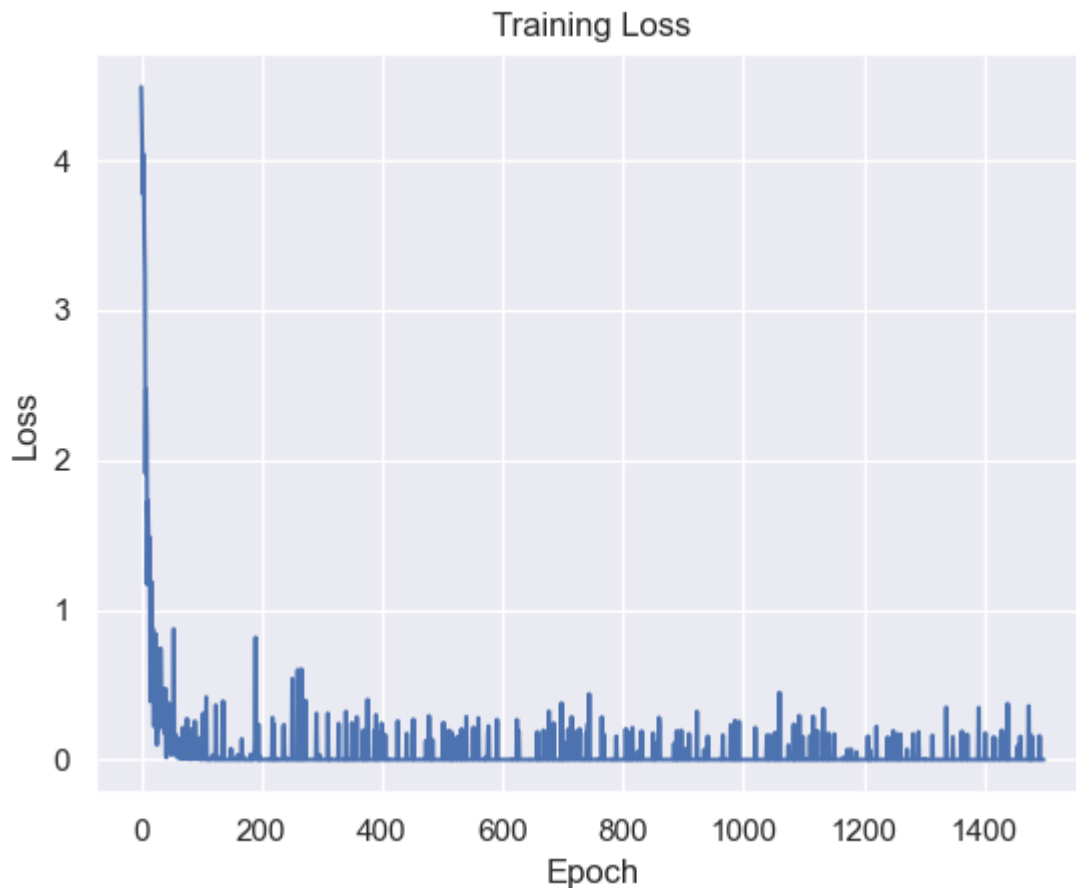
        #Calculate accuracy of model
        predicted_labels = torch.argmax(outputs, dim=1)
        correct_predictions +=(predicted_labels ==labels).sum().item()
        all_predicted_labels.extend(predicted_labels.tolist())
        all_true_labels.extend(labels.tolist())
    if (epoch+1) % 100 == 0:
        print (f'Epoch [{epoch+1}/{num_epochs}], Loss: {loss.item():.4f}')
    loss_values.append(loss.item())

```

*Listing 1.1: Sample code for unit testing*

The hyperparameter finetuning in this unit testing in *Listing 1.1* above sample code involved adjusting variables like `num_epoch`, `batch_size`, `learning_rate`, and `hidden_size` to analyze the behavior, performance, and accuracy of the model while undertaking training; The evaluation metrics, including classification reports, confusion matrix, determined the model's performance in this unit testing of the GRUs RNN model. The loss trend indicated if the model effectively learned data patterns as the weights and biases are been adjusted during the training. Validation data were used to test the model's predictive capabilities to validate the models' classification capabilities when deployed. This iterative unit testing process in the model training continued until optimal results were achieved based on loss trends,

classification reports, and validation performance. Different testing approaches were applied to the model training in a concurrent "test on the go" fashion, combining building and testing. *Figure 4.1* below represents the training loss of the model with the following parameters; num\_epoch=1500, batch\_size=10, learning\_rate=0.0012, input\_size=len(X\_train[0]), hidden\_size=8, and output\_size=len(tags) which achieved an accuracy of 98% and 0.000 final loss.



*Figure 4.1: Training Loss Trend of the Chatbot Model*

#### **4.2.2 Integration Testing**

After the complete implementation of all unit testing and achieving satisfactory results of various components of the chatbot solution, the system integration is done by putting all components together to function as a single unit. Integration testing is conducted to ensure all system components interact and function effectively as planned, it also helps to test the robustness of the system and the reliability of each of the system components. For example, using the integrated chatbot solution, the following components were tested:

- i. Hosting of the chatbot on local machine and visualizing the UI of the app on a web browser.
- ii. Verifying the accuracy of end users' queries and obtaining corresponding responses from the model predictions via the GUI.

- iii. Verifying the effectiveness of language APIs in detecting and translating queries and corresponding responses to users.
- iv. Evaluating the speed and accessibility of the system in processing users queries and requests.
- v. Assessing the overall performance of the Multilingual AI Chatbot in its engagement with users.

Thus, integration testing ensures the seamless collaboration of all components in the application developed in a reliable and efficient software solution in achieving its targeted goals. Therefore, integration testing helps verifies if the interconnected interfaces/components operate in a specific order, with each subsequent interface/component depending on its preceding counterpart. When each interface/component is invoked, such as "prompt for query entry," a check is performed on the language selection and the language API's capability to detect, translate, and process queries and corresponding responses; Queries and responses are passed to the next interface for its respective operations which are abstracted with the help of the UI. These interface checks facilitate the identification of hidden errors, enabling precise tracing of the specific unit responsible for any malfunctions. Integration testing goes beyond data-related errors and encompasses validating command transfer between interfaces and the communication and execution order of multiple processes in the chatbot system.

### **4.2.3 High-level Testing**

High level testing is the final phase of testing in this study's software implementation. The complete application was subjected to real-time scenarios, by conducting a real-time end-users/ clients' interactions with the chatbot. The objective of this high-level testing is to observe and verify the system's behavior and overall system performance in real-time. An example of how this was applied in this project is as stated below:

Once all project components have been completed and integrated into a unified system, a demonstration of the system's actual usage was conducted, and the execution time of each process was carefully observed to identify any unnecessary delays in the task completion; such as query processing, language detection and translation, and accuracy of each response as predicted by the system model. If delays or any kind of lapses were identified in this process, the project units associated with those tasks were examined by debugging and improving the unit to enhance overall system efficiency. Additionally, aspects such as the system's visual appearance, information presentation, and user-friendliness were critically evaluated during runtime, and appropriate improvements were made as necessary to improve total users' friendliness and satisfaction.

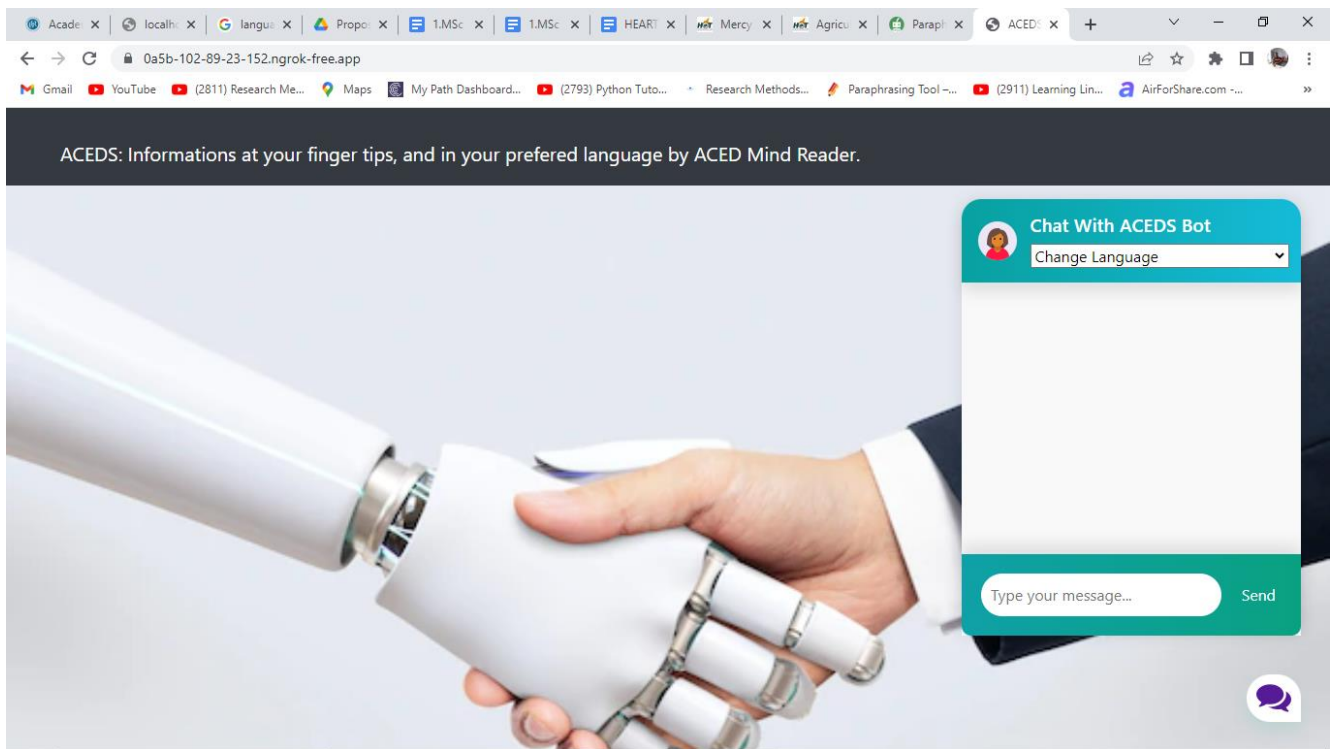
### **4.3 Result Discussions and Evaluation**

This phase of the study discusses the results of the study, and the evaluation of the overall system performance. The chatbot system comprises of various components that work together as a single system, and these components include the user interface, the language detection and translation component, the model training and validation components, the model for prediction of users' queries. The evaluation metrics of the system present the system performance. The study metrics include the confidence level, classification reports, and confusion matrices. These metrics were all printed to the system log during training, validation and real-time testing of the chatbot. The characteristics and outcome of the user interface are described below:

#### **4.3.1 GUI, and User's Queries/Response Documentation (Conversing with the Chatbot:)**

During the implementation/production phase, the application was hosted and tested on a local server on the personal computer through port 5000 (<http://localhost:5000>). Furthermore, a demo version of the application is made publicly available by using NGROK token to make the local host accessible on the internet for the real time demo testing.

Upon hosting and accessing the application, the (GUI) is displayed and includes a chatbot icon located at the bottom right corner of the web app welcome page. The default targets response language for consultation on the bot is English. Users can either use the default settings to converse and receive all responses in English or can converse in multiple languages and still receive responses back in English. Users can as well select a different specific target language for responses while performing multi-lingual queries on the bot. This functionality essentially allows users to engage in conversations in their preferred language and receive responses accordingly. The chatbot application when launched is visually represented in Figure 4.2 below.



*Figure 4.2: Chatbot Interface with Chosen English Response language*

The Chatbot system UI is as shown in *Figure 4.2 above*, the chatbot interface allows users to select their preferred consultation language of which users are much more comfortable with. In the figure above, the user is on default language which is English. By default, the chatbot respond to all multi-lingual queries in English. That is, the system has the capability to automatically detect the language of the conversation and provide responses back to users in English language. These unit functionality allows flexibility to switch to other languages such as Kirundi, Kinyarwanda, Hausa, French, Luganda, and so on. *Figure 4.3 below* is an illustration of the multi-lingual conversation while the users target response language is Kinyarwanda. In this scenario, the user has the leverage to engage the system with conversations in Kinyarwanda or any other languages available on the solution while receiving responses in Kinyarwanda.

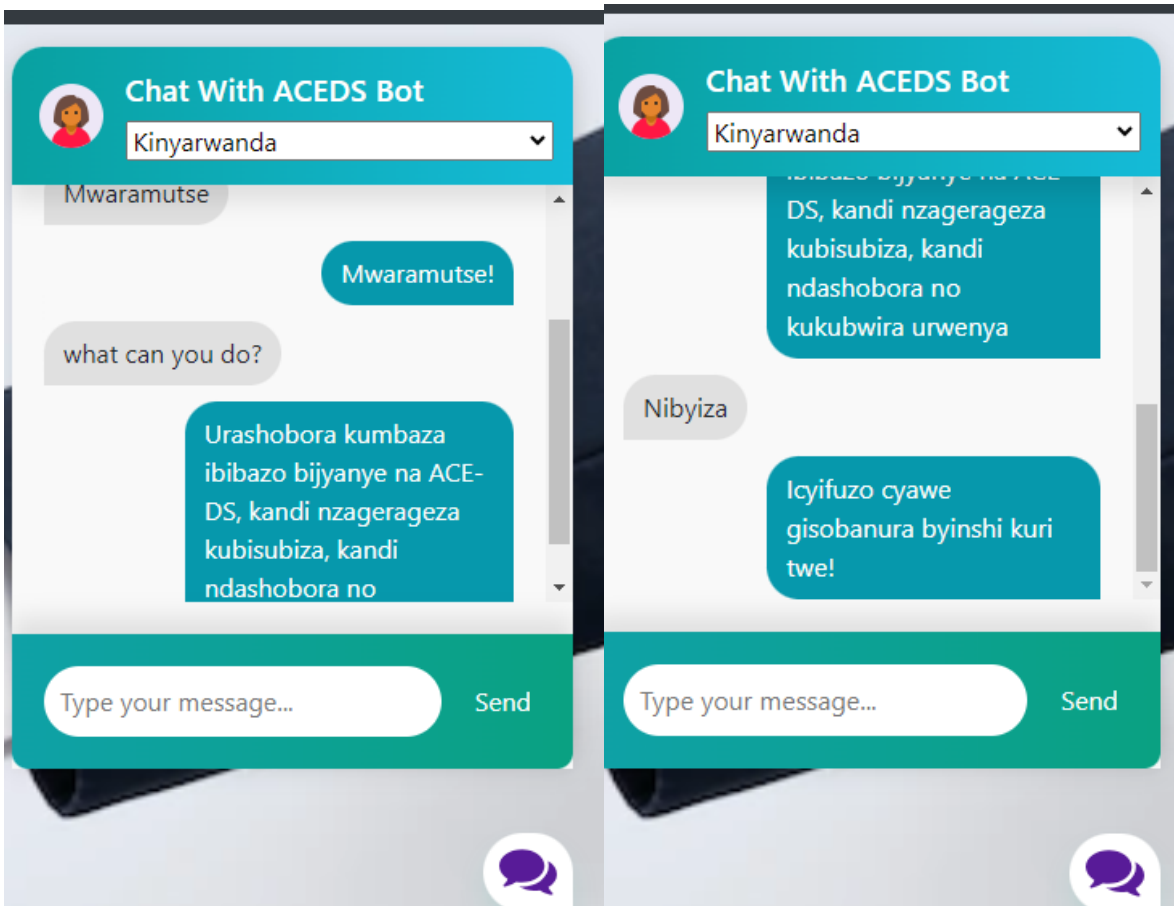


Figure 4.3: Multi-lingual Conversation with the chatbot system

### 4.3.2 Model Evaluation on the Training Data

The GRUs RNN deep learning model of the Chatbot was evaluated during the model training using the training data, for the purpose to understand how well the data was fitted in the model and the model performance in learning the hidden patterns in the dataset(intents). The training evaluation process of the neural network model uses Confusion Matrix and Classification Reports metrics imported from scikit learn machine learning python package. The evaluation is presented in Matrix format for the Confusion Matrix, Tabular format for the Classification report and the graphical loss trends showcasing the loss as the model adjust its weights and biases. Table 4.1 below is the classification report detailing the evaluation of the model on the training data.

**Table 4.1: GRUs RNN Model Performance Classification Report on Training Dataset**

	Precision	Recall	f1-score	Support
AI	1.00	0.98	0.99	15000
About	0.97	0.93	0.95	28500
Bachelor	0.93	0.91	0.92	16500
Certifications	0.92	0.95	0.93	18000
Degrees	1.00	0.98	0.99	16500
English	0.99	0.99	0.99	16500

GPA	0.99	0.97	0.98	16500
Grades	0.96	0.95	0.96	28500
Jobs	1.00	0.99	0.99	16500
Masters	1.00	0.99	0.99	13500
PHD	0.99	0.99	0.99	19500
PhD_Enrolment	1.00	0.97	0.99	15000
Reasons	0.99	0.99	0.99	25500
Transportation	1.00	0.99	0.99	21000
VISA	1.00	0.99	0.99	16500
admission	0.97	0.99	0.98	43500
advise	1.00	0.99	0.99	16500
anxious	1.00	0.98	0.99	13500
calender	0.99	0.99	0.99	18000
canteen	1.00	0.98	0.99	33000
certifications	0.99	0.98	0.99	18000
changes	0.98	0.99	0.98	24000
committee	1.00	0.98	0.99	12000
commuting	1.00	0.98	0.99	16500
content	1.00	0.99	0.99	16500
course	0.93	0.94	0.94	49500
creator	1.00	0.99	0.99	27000
credit	0.90	0.91	0.91	15000
deadline	1.00	0.99	1.00	16500
document	1.00	0.98	0.99	21000
establishment	1.00	0.98	0.99	21000
event	1.00	0.98	0.99	18000
excited	1.00	0.97	0.99	13500
facilities	0.85	0.87	0.86	15000
fees	0.99	0.99	0.99	52500
floors	1.00	0.98	0.99	15000
food	0.99	0.99	0.99	24000
goodbye	0.99	0.99	0.99	31500
greeting	1.00	0.98	0.99	25500
happy	1.00	0.97	0.99	18000
hostel	0.91	0.97	0.94	45000
hours	0.97	0.97	0.97	52500
housing	0.95	0.92	0.93	22500
infrastructure	0.79	0.73	0.76	9000
intake	1.00	0.98	0.99	24000
international	1.00	0.99	0.99	16500
jokes	0.99	0.99	0.99	45000
learning	1.00	0.99	0.99	18000
library	0.98	0.99	0.99	25500
living	1.00	0.98	0.99	16500
location	1.00	0.99	0.99	34500
master	0.93	0.90	0.91	16500
meaning	0.96	0.97	0.96	34500
menu	1.00	0.98	0.99	18000
modules	1.00	0.98	0.99	10500
name	1.00	0.98	0.99	28500
negative	1.00	0.99	0.99	19500
neutral	1.00	0.98	0.99	16500
number	0.99	0.99	0.99	28500
placement	0.99	0.99	0.99	21000
positive	0.98	0.99	0.98	45000
possibilities	0.97	0.99	0.98	16500
principal	0.93	0.91	0.92	18000
professional	0.99	0.99	0.99	13500
ragging	0.99	0.98	0.99	15000

random	1.00	0.97	0.98	6000
regulations	1.00	0.98	0.99	18000
sad	0.97	0.97	0.97	48000
safety	1.00	0.96	0.98	16500
salutaion	0.98	0.99	0.98	45000
scholarship	0.92	1.00	0.96	81000
semester	1.00	0.99	0.99	21000
speaking	0.99	0.98	0.99	16500
sports	1.00	0.99	0.99	15000
sports clubs	1.00	0.98	0.99	18000
staff	0.95	0.97	0.96	60000
state	1.00	0.98	0.99	16500
stressed	1.00	0.98	0.99	19500
swear	0.92	0.89	0.91	13500
switching	0.99	0.96	0.98	16500
syllabus	0.99	0.99	0.99	22500
task	0.98	0.99	0.99	33000
timeline	1.00	0.99	1.00	16500
transfer	1.00	0.99	0.99	16500
uniform	1.00	0.99	0.99	13500
vacation	1.00	0.99	0.99	19500
why	0.99	0.98	0.98	16500
work	0.98	0.98	0.98	16500
<b>accuracy</b>			<b>0.98</b>	<b>2034000</b>
<b>macro avg</b>	<b>0.98</b>	<b>0.97</b>	<b>0.98</b>	<b>2034000</b>
<b>weighted avg</b>	<b>0.98</b>	<b>0.98</b>	<b>0.98</b>	<b>2034000</b>

Table 4.1 above depicts the model training evaluation classification report. The evaluation report on the training data demonstrates excellent accuracy, precision, recall, and f1-score, as well as the weighted average of each metrics. The model achieved an impressive accuracy of 98%, while maintaining an average weighted precision, recall, and f1-score of 98% for classifying all the tags or classes in the training data. These metrics indicate that the model has effectively learned the data and can make accurate predictions. The snapshot in figure 4.4 below showcases the model's confusion matrix, highlighting the correct predictions located on the diagonal of the evaluation matrix.

```

Confusion Matrix:
[[14853  0  0 ...  0  0  0]
 [  0 26792  0 ...  0  0  0]
 [  0  0 15107 ...  0  0  0]
 ...
 [  0  0  0 ... 19343  0  0]
 [  0  0  0 ...  0 16325  0]
 [  0  0  0 ...  0  0 16382]]

```

Figure 4.4: Confusion Matrix of the model on the training data

### 4.3.3 Model Evaluation Using Validation Data

The system model was also validated using the validation data, and during the validation phase, the model underwent validation using unseen data to assess its performance and behavior with data it was not trained with. The primary objective of this validation check was to evaluate the model's ability to make accurate predictions and seamlessly integrate it with other components of the chatbot if its validation results shows the model wasn't overfitted or underfitted. In the validation process, evaluation metrics from Scikit-learn machine learning python package were employed to generate a comprehensive classification report detailing the accuracy, precision, f1-score, recall, macro average, and weighted average. The classification report of the validation provides a comprehensive evaluation of the model's predictions on unseen data. The results of this evaluation are presented in the table depicted in Table 4.2.

**Table 4.2: GRUs Chatbot Model's Classification Performance Reports on Validation Data**

Class	Precision	Recall	f1-score	Support
AI	1.00	1.00	1.00	10
About	1.00	0.95	0.97	19
Bachelor	1.00	0.91	0.95	11
Certifications	0.92	1.00	0.96	12
Degrees	1.00	1.00	1.00	11
English	1.00	1.00	1.00	11
GPA	1.00	1.00	1.00	11
Grades	0.95	1.00	0.97	19
Jobs	1.00	1.00	1.00	11
Masters	1.00	1.00	1.00	9
PHD	1.00	1.00	1.00	13
PhD_Enrolment	1.00	1.00	1.00	10
Reasons	1.00	1.00	1.00	17
Transportation	1.00	1.00	1.00	14
VISA	1.00	1.00	1.00	11
admission	1.00	1.00	1.00	29
advise	1.00	1.00	1.00	11
anxious	1.00	1.00	1.00	9
calender	1.00	1.00	1.00	12
canteen	1.00	1.00	1.00	22
certifications	1.00	1.00	1.00	12
changes	1.00	1.00	1.00	16
committee	1.00	1.00	1.00	8
commuting	1.00	1.00	1.00	11
content	1.00	1.00	1.00	11
course	1.00	0.94	0.97	33
creator	1.00	1.00	1.00	18
credit	1.00	0.90	0.95	10
deadline	1.00	1.00	1.00	11
document	1.00	1.00	1.00	14
establishment	1.00	1.00	1.00	14
event	1.00	1.00	1.00	12
excited	1.00	1.00	1.00	9
facilities	0.83	1.00	0.91	10
fees	1.00	1.00	1.00	35

floors	1.00	1.00	1.00	10
food	1.00	1.00	1.00	16
goodbye	1.00	1.00	1.00	21
greeting	1.00	1.00	1.00	17
happy	1.00	1.00	1.00	12
hostel	0.97	1.00	0.98	30
hours	0.97	1.00	0.99	35
housing	1.00	0.93	0.97	15
infrastructure	1.00	0.67	0.80	6
intake	1.00	1.00	1.00	16
international	1.00	1.00	1.00	11
jokes	1.00	1.00	1.00	30
learning	1.00	1.00	1.00	12
library	1.00	1.00	1.00	17
living	1.00	1.00	1.00	11
location	1.00	1.00	1.00	23
master	0.92	1.00	0.96	11
meaning	0.96	1.00	0.98	23
menu	1.00	1.00	1.00	12
modules	1.00	1.00	1.00	7
name	1.00	1.00	1.00	19
negative	1.00	1.00	1.00	13
neutral	1.00	1.00	1.00	11
number	1.00	1.00	1.00	19
placement	1.00	1.00	1.00	14
positive	1.00	1.00	1.00	30
possibilities	1.00	1.00	1.00	11
principal	1.00	0.92	0.96	12
professional	1.00	1.00	1.00	9
ragging	1.00	1.00	1.00	10
random	1.00	1.00	1.00	4
regulations	1.00	1.00	1.00	12
sad	0.97	1.00	0.98	32
safety	1.00	1.00	1.00	11
salutation	1.00	1.00	1.00	30
scholarship	1.00	1.00	1.00	54
semester	1.00	1.00	1.00	14
speaking	1.00	1.00	1.00	11
sports	1.00	1.00	1.00	10
sports clubs	1.00	1.00	1.00	12
staff	0.98	1.00	0.99	40
state	1.00	1.00	1.00	11
stressed	1.00	1.00	1.00	13
swear	1.00	0.89	0.94	9
switching	1.00	1.00	1.00	11
syllabus	1.00	1.00	1.00	15
task	1.00	1.00	1.00	22
timeline	1.00	1.00	1.00	11
transfer	1.00	1.00	1.00	11
uniform	1.00	1.00	1.00	9
vacation	1.00	1.00	1.00	13
why	1.00	1.00	1.00	11
work	1.00	1.00	1.00	11
<b>accuracy</b>			<b>0.99</b>	<b>1356</b>
<b>macro avg</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>1356</b>
<b>weighted avg</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>1356</b>

The table 4.2 above, is an illustration of the evaluation results of the model on unseen data, showcasing outstanding accuracy score, precision score, recall, f1-score and weighted average of the model on the data. The model achieved an impressive accuracy of 99% and also maintained an average precision, recall, and f1-score of 99% for its classification of all the tags in the validation intent. This validation results confirms and demonstrate that the model can effectively make accurate predictions based on its knowledge scope.

#### **4.3.4 Model Evaluation Through Real Time Testing**

In the real time testing of the app, the Multilingual AI Chatbot system underwent real-time test using users' Real Time queries by inputting consultation messages to verify if the various component of the system conforms with the study objective. This is done by hosting the system via local host and the internet with the help of NGROK putting the locally hosted app to the public domain. The purpose of this Realtime evaluation of the system is to assess its performance as a deployed web application demo. The real-time evaluation testing process also employed Scikit-learn machine learning python package to measure the confidence level and accuracy of the model in its prediction of users' real time queries/consultation. Thus, the objective of the real-time evaluation of the multi-lingual chatbot application include the following:

- i. Evaluation of the Chatbot responsiveness: The system real-time evaluation enables confirmation that the system adheres to the established reaction time specifications of the system and generates timely results to users.
- ii. Data integrity Validation: This is to ensure that the accuracy and integrity of data processing of the chatbot during the real time testing is optimal, that is, how accurately these data are been processed and returned to the end users in real time.
- iii. Evaluation of the chatbot reliability: The majority of the time, real-time systems are up and running continually without fatigue. Real-time evaluation of the Chatbot makes it possible to easily spot potential system flaws, delays, or errors that may occur and may have a chance to negatively affect its reliability. This helps to ensure that the system operates consistently under a range of different circumstances or conditions.
- iv. Functional requirement verification of the application: The real-time evaluation of the chatbot system confirms the system's compliance with the functional requirements thereby making sure it completes the intended tasks aimed for accurately and in real-time settings.

From the aforementioned, the researcher observed and document the real time evaluation of the chatbot solution. The researcher also observed the speed of the system in processing users' consultations in cross-language communication. Table 4.3 below presents an evaluation documentation of this real-time testing of the chatbot.

**Table 4.3: Real Time Chatbot Testing/Evaluation**

<b>Input messages or Query</b>	<b>Intent or tag</b>	<b>Intent Prediction</b>	<b>Confidence Level (0-1)</b>
hello dear	greeting	Greeting	1
what can you do?	Tasks	Tasks	0.999957
Does ACEDS have state of the art facilities for learning?	facilities	facilities	0.999747
How much is ACEDS fees?	Fees	Fees	1
How can i apply for admission at ACEDS?	admission	admission	1
What master's programs does ACEDS offers?	Master or Programs	programs	0.999387
What are the necessary academic requirements for enrolling in the Masters program?	master	master	1
information about masters programs	master	master	0.999966
Tell me about ACEDS?	About	About	0.999742

Table 4.3 above presents a sampled real-time testing report of the chatbot system, the report shows that the system can effectively make accurate predictions confidently based on its knowledge scope. Hence, it can be concluded that the solution demonstrates excellent real-time predictions of user engagement with above 95% confidence level when presented with queries related to the trained classes or available information in its knowledge scope. However, in the cases where the model encounters a query outside its knowledge scope or topics, a fallback prompt is utilized as a response to the user. It is important to acknowledge that there are areas that need further improvement, which include the implementation of the attention mechanism or the transformer model, incorporation of reinforcement learning, and or the use of transfer learning algorithms which have the potential for better performance and generalization. Additionally, the model needs to be trained with a larger quality dataset to achieve an advanced sophistication expected to cover a wider educational consultation scope.

### 4.3.5 User Satisfaction Evaluation Through Users Testing

Users evaluation testing was conducted with survey taken to get feedbacks from users for the purpose of ascertaining the performance of the study solution; This survey intends to gather feedback from users regarding their experience with the chatbot solution in ACEDS information provisioning. The analysis and the description in figure 4.5 below detailed the surveys feedback from Fourteen (14) respondents expressing their experience using the research solution;

**Table 4.4: User Satisfaction Survey Evaluation of the Chatbot Solution**

Questions	Response 1	Response 2	Response 3	Response 4	Response 5
How do you rate the Vocabulary, mechanical accuracy, sentence construction and soundness of expression of the responses provided by the Chatbot?	(0-30%) =0% users	(31-50%) =7.1% users	(51-80%) =50% users	(81-100%) =42.9% users	
Does the system converse efficiently using your preferred language (e.g English, French, Kinyarwanda etc)?	Yes=71.4%	No=0%	Partially=28.6%		
How satisfied are you generally with the answers provided by the cha bot in terms of efficiency and accuracy?	(0-30%) =7.1% users	(31-50%) =14.3% users	(51-80%) =57.1% users	(81-100%) =21.4% users	
In which of the following categories did the chatbot perform poorly in terms of the correctness and usefulness of the answers provided?	Academic-related questions =21.4% users	Accommodati on related questions =21.4% users	Admission enquiry =21.4% users	History about the institution =21.4% users	Socializing and Extracurricular =7.3% users
Did the chatbot answer your questions promptly and correctly?	Yes=64.3%	Partially =28.6%	No=7.1%		
Which features are most valuable to you?	The promptness of response provided=35.7 %	The accuracy of responses provided =35.7%	The interactivity, choice of design and display of output =28.6%		
Did the chatbot software contribute to a more efficient and time-saving inquiry session for you?	Yes=92.9%	No=7.1%			
How well did the chatbot software adapt to your individual needs and learning style within the educational context?	Perfectly =35.7%	Average =57.1%	Poorly =7.1%		
Overall, how likely are you to recommend this chatbot software to other students or educators in an educational setting?	Certainly =28.6%	Highly unlikely =71.4%	Unlikely =0.0%		

The survey in Table 3.4 above indicated that 71.4% of users are satisfied with the efficient use of their preferred language in consultations. The chatbot performed poorly in terms of responding to queries that are historical and descriptions about the institution which was not part of the training dataset. Queries related to academics and accommodation had the best ratings from

users. In terms of the promptness of responses provided by the chatbot, 64.3% of users in the survey indicate satisfaction, and in terms of more efficient time-saving sessions and ease of information acquisition, the feedback indicates 92.9% satisfaction. In terms of the features of the chatbot which users find most valuable, the survey indicated the high accuracy of the responses provided by the chatbot is the most valuable as opposed to design and interactivity features.

#### **4.3.6 Users Recommendation**

**Below are some of the features users recommended to improve the system performance**

*What important features do you think are missing in the Chatbot?*

- i. The voice note feature, which can be useful for visually impaired persons to help them interact with the chat bot.
- ii. Auto correct and suggestions or assistance with helping the completion questions been typed by clients
- iii. The chatbots doesn't really have all information; Hosted online, Design
- iv. More understanding of languages
- v. Prompt user to contact person in charge for inability to respond
- vi. You need to train the chatbot with more data for better performance

## CHAPTER FIVE: CONCLUSION AND RECOMMENDATIONS

### 5.1 Conclusion

The chatbot developed in this research, is a web-based software solution specifically designed as an artificial consultation and information provisioning agent to cater to the information needs of students, general public, staff, and prospective students of ACEDS, University of Rwanda. The research demonstrated that this solution can efficiently and accurately provide users with essential information they may struggle to obtain directly from ACEDS management with a high level of precision and accuracy. A key contribution demonstrated by this study is the use of between 1 to over 50 languages for conversing between users and the bot where the dataset used for training was collected in only one language (English) rather than using the conventional approach where datasets must be curated in all languages used. This is a great improvement in terms of efficiency because the model is not trained to learn hundreds of languages which is time consuming but rather, queries and responses are analyzed in English and translators are used and responses translated to the clients preferred language of conversing. This is a novel approach to resolving AI conversational challenges that require interpretations in many languages.

The results accrued from the chatbot usage were documented and presented in a tabulated and visualized format, highlighting the chatbot's prediction capabilities in various scenarios. The chatbot machine learning model achieved a high accuracy of 98% in classifying seen data. Whereas, in the validation phase, the system achieved an accuracy of 99% on the validation data. Also, in the real-time testing, the integrated system consistently achieved an accuracy and confidence level of accurate predictions of above 95% of its knowledge-based information. The chatbot performance demonstrated a reliable accuracy in its ability to make predictions, detect a language and communicate across different languages, effectively there by achieving the desired multilingual capability of the research aims and objective. Thus, the overall performance of the model indicates that it successfully achieved the objectives of the study. However, it is important to note that the prediction scope of the model is limited due to insufficient training data.

On a general note, the chatbot performed satisfactorily as the average accuracy was above 95%. Additionally, the user satisfaction survey indicated that over 50% of the correspondents are satisfy with the chatbot performance in correctness of grammar usage, contribution to a more efficient and time saving, efficient use of preferred language, responses provided, prompt and correct response provisioning, and time saving in information acquisitions. The survey further indicated that the chatbot performed poorly in terms of responding to queries that are historical and descriptions about the

institution. Queries related to academics and accommodation had the best ratings from users. In terms of the promptness of responses provided by the chatbot, 64.3% of users in the survey indicate satisfaction, and in terms of more efficient time saving sessions and ease of information acquisition the feedback indicates 92.9% satisfaction. In terms of the features of the chatbot which users find most valuable, the survey indicated the high accuracy of the responses provided by the chatbot is the most valuable as opposed to design and interactivity features.

## **5.2 Recommendations**

Some areas for improvement that should be addressed in order to achieve a more sophisticated solution expected for a standard software in future study are as follows:

From the user satisfaction survey, 1 respondent recommended the use of the voice note feature, which can be useful for visually impaired persons to help them interact with the chat bot. The researcher deems this recommendation important because it will make the chatbot inclusive for visually impaired persons and amputees who cannot use their hands for typing their queries. Hence, a suggestion on how to accomplish this is the use of Google speech to text API (<https://cloud.google.com/speech-to-text>) One(1) respondent, reiterated the need for autocorrect and suggestions/assistance as a prompt in the chat text field which will help the completion questions been typed by clients. This feature can be incorporated using open source APIs typically available from GitHub([www.github.com](http://www.github.com)).

Furthermore, the quantity and quality of the training datasets used for the model need enhancement and also large enough for the model fitting. According to Goodfellow, Bengio, and Courville( 2016) limited data and information can lead to underfitting or overfitting. Also, fallbacks may occur when the model cannot find sufficient information in user input messages to make accurate classifications due to limited training data. Hence, the chatbot solution provided had limited vocabulary which is a function of insufficient dataset used for training the model because the learning and brilliance of the models were dependent on the scope and availability of data, hence when questions were asked beyond the data scope, the model found it difficult to respond with precise and pertinent information. The sentiment data was particularly insufficient hence, the chatbot did not properly comprehend the feelings expressed by users during conversing. It occasionally replied insensitively in sensitive or sympathetic circumstances. This challenge impeded the attainment of one of the predefined objectives of this research i.e. developing an emotionally intelligent chatbot. A possible way to enhance the volume, variation and quality of the training data set is the use of data augmentation.

Furthermore, the chatbot solution was unable to keep track of conversations i.e. It attends to each query as an entirely new query whereas in some circumstances, the queries presented by users are leading and follow-up questions to the previous conversations.

The implementation of large language models such as transformer models, transfer learning, or reinforcement learning techniques can enhance the overall system performance in predicting user responses to queries. This language models can be improved via fine-tuning, data filtering. Furthermore, future research should focus on implementing technologies that enable efficient documentation and capturing of all data exchanged during client-system interactions, as this record are essential for further analysis and improvement of the software. This would involve the integration of a structured and efficient database system to store user data and conversations. Additionally, the implementation of reinforcement learning technique to facilitate continuous learning by the model, using user input messages and interaction records to retrain itself and improve its information provisioning capacity and prediction accuracy over time.

## References:

- Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, 31(2), 427–445. <https://doi.org/10.1007/s12525-020-00414-7>
- Alruqi, T. N. and Alzahrani, S. M.(2023), "Evaluation of an Arabic Chatbot Based on Extractive Question-Answering Transfer Learning and Language Transformers," *Ai*, vol. 4, no. 3, pp. 667–691, 2023, doi: 10.3390/ai4030035.
- Ahuja, K., Hada, R., Ochieng, M., Jain, P., Diddee, H., Maina, S., ... & Sitaram, S. (2023). Mega: Multilingual evaluation of generative AI *arXiv preprint arXiv:2303.12528*.
- Anand, S., Karthikeya, M., Sai, A. A., & Balamurali, O. (2023, January). Multi-Lingual Hybrid Chatbot for Empowering Rural Women Self-Help Groups in India. In *2023 International Conference for Advancement in Technology (ICONAT)* (pp. 1-6). IEEE.
- Aslani, S., & Jacob, J. (2023). Utilisation of deep learning for COVID-19 diagnosis. *Clinical Radiology*, 78(2), 150-157.
- Bharti, U., Bajaj, D., Batra, H., Lalit, S., Lalit, S., & Gangwani, A. (2020, June). Medbot: Conversational artificial intelligence powered chatbot for delivering tele-health after covid-19. In *2020 5th International Conference on Communication and Electronics Systems (ICCES)* (pp. 870-875). IEEE. <https://ieeexplore.ieee.org/abstract/document/9137944>
- Biswas, D. (2020). Privacy Preserving Chatbot Conversations. *Proceedings - 2020 IEEE 3rd International Conference on Artificial Intelligence and Knowledge Engineering, AIKE 2020*, 179–182. <https://doi.org/10.1109/AIKE48582.2020.00035>
- Brouwer, R. J., Vasterling, J. J., Vermeiren, R., & Schmand, B. (2021). The impact of artificial intelligence on work and employment: A systematic review of the literature. *Journal of Business Economics*, 91(1), 189-224.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Goodfellow, I. (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.
- Chung, J. H., Kim, H. S., Kim, Y. H., & Kim, K. H. (2021). The impact of chatbots on healthcare services: A systematic review. *Journal of Medical Systems*, 45(6), 369.
- Data Science Process Alliance. (2023). CRISP-DM Diagram. <https://www.datascience-pm.com/crisp-dm-2/>
- Davison, R. M., Martinsons, M. G., and Kock, N. (2004). Principles of Canonical Action Research. *Information Systems Journal* 14(1), 65-86.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Dickens, L., & Watkins, K. (1999). Action research: rethinking Lewin. *Management Learning*, 30(2), 127-140. doi: 10.1177/1350507699302002
- Evgeniia, R., Goran, G., Olga, M., Anna, K., and Ivan, V. *Crossing the Conversational Chasm: A Primer on Multilingual Task-Oriented Dialogue Systems*.
- Firdaus, M., Ekbal, A., & Bhattacharyya, P. (2020, May). Incorporating politeness across languages in customer care responses: Towards building a multi-lingual empathetic dialogue agent. In *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 4172-4182).

<https://aclanthology.org/2020.lrec-1.514/>

- Floridi, L. (2023). AI as Agency without Intelligence: On ChatGPT, large language models, and other generative models. *Philosophy & Technology*, 36(1), 15.
- Gaurav, K. and Pradeep, K.B., (2012). Impact of Agile Methodology on Software Development Process. *International Journal of Computer Technology and Electronics Engineering(IJCTEE)*, 2(4), pp. 1-5.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
- Kanwisher, N., Khosla, M., & Dobs, K. (2023). Using artificial neural networks to ask ‘why’ questions of minds and brains. *Trends in Neurosciences*.
- Khaza, N.M. and Syam, C.K., 2020. An analytical approach in usage of agile methodologies in construction industries – A case study. *Materialstoday: proceedings*, 33(1), pp. 475-479
- Koehler, J., Fux, E., Herzog, F. A., Lötscher, D., Waelti, K., Imoberdorf, R., & Budke, D. (2018). Towards intelligent process support for customer service desks: Extracting problem descriptions from noisy and multi-lingual texts. In *Business Process Management Workshops: BPM 2017 International Workshops, Barcelona, Spain, September 10-11, 2017, Revised Papers 15* (pp. 36-52). Springer International Publishing.  
[https://link.springer.com/chapter/10.1007/978-3-319-74030-0\\_3](https://link.springer.com/chapter/10.1007/978-3-319-74030-0_3)
- Kohavi, R., & Provost, F. (1998). special issue on applications of machine learning and knowledge discovery. *encyclopedia of machine learning*, 271-274.
- Korthikanti, V. A., Casper, J., Lym, S., McAfee, L., Andersch, M., Shoeybi, M., & Catanzaro, B. (2023). Reducing activation recomputation in large transformer models. *Proceedings of Machine Learning and Systems*, 5.
- Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. *Annals of Data Science*, 10(1), 183-208.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems 25* (pp. 1097–1105).
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- Lin, C. C., Huang, A. Y., & Yang, S. J. (2023). A review of ai-driven conversational chatbots implementation methodologies and challenges (1999–2022). *Sustainability*, 15(5), 4012.
- Mahboubeh, D., Saeed, A. and Samad, P., (2020). *A Systemic Literature Review on Semantic Web-enabled Software Testing*. *Journal of Systems and Software*, 162 pp. 110485
- Mattas, P. S. (2023). ChatGPT: A Study of AI Language Processing and its Implications. *Journal homepage: www.ijrpr.com ISSN, 2582, 7421*.
- Mazhar F.(2020). Multilingual Chatbots (LUIS & Watson). Government College University, Lahore.
- Mehrish, A., Majumder, N., Bharadwaj, R., Mihalcea, R., & Poria, S. (2023). A review of deep learning techniques for speech processing. *Information Fusion*, 101869.
- Michael, R., Christian, D., Stefan, P., Philipp, L., & Guenther, S. (2021). *Methodology for iterative system modeling in agile product development*. *Procedia*, 100, pp. 439-444.
- Min, E., Chen, R., Bian, Y., Xu, T., Zhao, K., Huang, W., ... & Rong, Y. (2022). Transformer for

graphs: An overview from architecture perspective. *arXiv preprint arXiv:2202.08455*.

- Müller, C., Kratzwald, S., & Neubauer, A. (2019). Chatbots in customer service—A systematic literature review. *Business & Information Systems Engineering*, 61(4), 367-378.
- Narayan, V., Mall, P. K., Alkhayyat, A., Abhishek, K., Kumar, S., & Pandey, P. (2023). Enhance-Net: An Approach to Boost the Performance of Deep Learning Model Based on Real-Time Medical Images. *Journal of Sensors*, 2023.
- Powles, J. (2020). *The ethics of artificial intelligence*. Cambridge University Press.
- Pradana, Z. H., & Rochmanto, R. A. (2022). *Chatbot-based Information Service using RASA Open-Source Framework*. 5(158), 2–3.
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why Should I Trust You?: Explaining the Predictions of Any Classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144).
- Razumovskaia, E., Glavaš, G., Majewska, O., Korhonen, A., & Vulic, I. (2021). Crossing the conversational chasm: A primer on multilingual task-oriented dialogue systems. *arXiv preprint arXiv:2104.08570*. [https://evgeniiaraz.github.io/files/Multilingual\\_ToD\\_Primer.pdf](https://evgeniiaraz.github.io/files/Multilingual_ToD_Primer.pdf)
- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems 27* (pp. 3104–3112).
- Tarun, L., Shashank, B., Ashish, P., Shreya, B., Vasundhara, R. (2018). Implementation of a Chatbot System using AI and NLP. *International Journal of Innovative Research in Computer Science & Technology (IJIRCST)*. ISSN: 2347-5552, Volume-6, Issue-3, May 2018. DOI: 10.21276/ijircst.2018.6.3.2.
- Vahid, G., Michael, V. and Feyza, N.K. (2019). A survey on software testability. *Information and Software Technology*, 108 pp. 35-64.
- Yang, L. W. Y., Ng, W. Y., Lei, X., Tan, S. C. Y., Wang, Z., Yan, M., ... & Ting, D. S. W. (2023). Development and testing of a multi-lingual Natural Language Processing-based deep learning system in 10 languages for COVID-19 pandemic crisis: A multi-center study. *Frontiers in public health*, 11, 387.
- Zhou, M. (2022). *Towards Multi-lingual Multi-modal Dialogue Systems*. University of California, Davis. <https://www.proquest.com/openview/c0976bffbbaa23a6c759294b929839d7/1?pq-origsite=gscholar&cbl=18750&diss=y>.
- ZXhang, Y. X., Haxo, Y. M., & Mat, Y. X. (2023). *Can Neural Networks Predict Stock Market?* (No. Stock Analysis). AC Investment Research.

## Appendix I

### KNOWLEDGE DATA INFORMATION GATHERING SAMPLE

SN	Questions	Response (Answers)	Information Links (If available)
1	What is the full meaning of ACEDS?	African Center of Excellence in Data Science(ACE-DS)	
2	How do I get enrolled at ACEDS	When the call is out, the African Centre of Excellence in Data Science welcomes applications from candidates with relevant undergraduate/master's degrees in Statistics, Mathematics, Computer Science, Information Technology, Economics and other disciplines related to Data science. If all requirements are fulfilled then you get enrolled	
3	What is the motto of ACEDS?		
4	What are the academic requirements needed to get enrolled for Masters programme?	Relevant undergraduate degree in Statistics, Mathematics, Computer Science, Information Technology, Economics and other disciplines related to Data science.	
5	"What are the academic requirements needed to get enrolled for PhD programme?"	Relevant master's degrees in Statistics, Mathematics, Computer Science, Information Technology, Economics and other disciplines related to Data science.	
6	"What are the academic requirements needed to get enrolled for Professional Certification programme?"	Having a bachelor's degree in the relevant field in Data Science	
7	What are the visa requirements for international students?	Upon arrival, students are granted a visa, and the necessary documents include a Police Clearance Certificate from their home country, a recommendation letter	

		from ACEDS, a Medical Certificate of Fitness, an International Passport, and passport photographs.	
8	What are the best things about education at ACEDS?	The availability of diverse courses, events, and international connections enhances your employability and job prospects.	
9	When is the application deadline or timeline and does it make a difference if I apply early?	Please kindly refer to our application announcements, and we encourage applicants to submit their applications well in advance or by the stated deadline.	
10	How do I get a job on campus?	On-campus employment opportunities for students are not available. However, exceptional students may be occasionally recruited for specific positions, which will be communicated to them in due course.	
12	what is public transport like whilst enrolled at ACEDS?	Rwanda offers a wide range of transportation options, including buses, metros, taxis, personal cars, bicycles, and more.	
13	How do I get involved in a student club or sport?	Participating in clubs or sports can be an enriching aspect of your ACEDS journey, and there are dedicated sporting facilities available to enhance this experience.	
14	Do you offer on-campus housing?	The availability of on-campus apartments for students is limited. However, there are various affordable accommodation options in the vicinity of the campus environment.	
15	Can I speak to someone from ACEDS in my native language?	Yes if the person understands your native language but kindly note that the language of instructions is English	
16	What are the costs of studying at ACEDS?	<a href="https://ur.ac.rw/IMG/pdf/fees_structure_2018_2019_combined-pivote.pdf">https://ur.ac.rw/IMG/pdf/fees_structure_2018_2019_combined-pivote.pdf</a>	<a href="https://ur.ac.rw/IMG/pdf/fees_structure_2018_2019_combined-pivote.pdf">https://ur.ac.rw/IMG/pdf/fees_structure_2018_2019_combined-pivote.pdf</a>
17	What's the food really like in kigali? Can I get food from my home country?	Rwanda offers a diverse range of food options, and as a country known for its tourist attractions, you can find inter-continental cuisine available as well.	
18	How can I get a scholarship at ACEDS?	Scholarships at ACEDS are granted through a competitive selection process. To explore the available scholarships, please refer to our application announcements.	
19	When does an academic calendar for ACEDS start?	The academic calendar begins once the admission process is completed, while for Professional programs, it starts immediately after enrollment.	

20	What are the different types of graduate degrees offered at ACEDS?	ACEDS offers postgraduate programs encompassing MSc and PhD Degrees in Data Science with few areas of specialization	
21	Is it possible to obtain a bachelor's degree and a master's degree at the same time?	No, ACEDS only offers postgraduates programs encompassing Masters and PhDs degrees programs in Data Science	
22	What is the <b>average</b> length of study for Masters programs at ACED?	Two years for Masters Programs (24 Months) Three years for PhD programs(36 months)	
23	Can you work while studying at ACEDS?	Yes, you can enrol for day or evening programs depending on which program is convenient for your work.	
24	What is the difference between online studies and distance studies?		
25	Is distance learning available at the graduate level?	ACEDS exclusively provides a blended learning system that combines both traditional and virtual methods of education.	
26	Can I transfer to a ACEDS from a university outside of the Rwanda?	No, but kindly contact the center admin for advice on how to get enroled	
27	Do students have to study a fixed set of courses or can they individually form their major?	Depending on their area of specialization, students are required to complete a specific number of courses.	
28	What are English language proficiency requirements?	Proficiency in both written and verbal communication is expected from students at ACEDS as all programs are instructed in English.	
29	What is a GPA?	GPA stands for Grade Point Average. It is a numerical representation of a student's academic performance..	
30	How does the credit system at ACEDS work?	The grading system adopted by ACEDS is approved by the Higher Education Council and the Senate of the University of Rwanda.	
31	How are grades determined in ACEDS?	Typically, classwork accounts for 60% of the grade, while exams contribute 40% for each module undertaken.	

32	Can an enrolled student change his/her assigned course of study?	Normally, program changes are not permitted; however, exceptions may be made if a valid reason is provided to the management. For further information, please reach out to the center administrator.	
33	What are the steps required for changing department/course?	To proceed with this matter, kindly get in touch with the center administrator.	
34	How can an international student find out what academic courses from their country are acceptable for at ACEDS?	International students can refer to our calls for applications to stay updated and, if needed, they can reach out to the contact information provided.	
35	How can I apply for professional certification training programs?	Contact the center for details on how to get enrolled	
36	What's the best general advice for an incoming student?	Build relationships with individuals who uphold strong social, academic, and personal principles. Remember that honesty is the most effective approach. Steer clear of all forms of cheating, including collusion and plagiarism.	
37	Why should I choose to study at ACEDS	ACEDS at the University of Rwanda has received international accreditation from the Data Science Council of America (DASCA). The center recruits students and lecturers from various countries.	
38	How much should I expect to spend as living cost for a year in Rwanda.	On average, a monthly expense, including student accommodation rent, can be comfortably covered by \$300 USD, but this may vary depending on individual lifestyle and spending patterns.	
39	How do I commute around Rwanda? What is the transport system like?	Getting around Kigali and other regions in Rwanda is a convenient, seamless, and secure experience due to the availability of various transportation systems.	
40	How safe is Rwanda or Kigali	Rwanda has long been considered one of the safest countries in Africa to travel to. This landlocked East African country is home to rare animals, making it a dream destination for nature lovers.	

41	what state of the Art Facilities are there at ACEDS?	ACEDS offers state-of-the-art facilities, including modern classrooms, well-equipped laboratories, a comprehensive library, and dedicated computer labs.	
46			

## Appendix II

### User Satisfaction Evaluation Questionnaire

#### User Satisfaction Evaluation Form in respect to the Chatbot solution provided as part of an MSc Thesis

This surveys intends to gather feedback from users regarding their experience with the chatbot solution (Case study of ACEDS Information provisioning)-ACEDS Consultant Agent. The chatbot was intend to answer queries encompassing programs information's, scholarships, syllabus, fees, registration processes, required documents for admission, accommodation, events, training facilities, sport and club, and so on. You can select a prefered language of communication and interact with the system in multiple languages and provide feedbacks.

Click this link to access the chatbot : <https://e79b-102-91-5-83.ngrok-free.app/>

Please provide your responses as regards your experience from interacting with the Chatbot

How do you rate the Vocabulary, mechanical accuracy, sentence construction and soundness of expression of the responses provided by the Chatbot?

- 0-30%
- 31-50%
- 51-80%
- 81-100%

Does the system converse efficiently using your preferred language(e.g English, French, Kindarwanda etc) ?

- Yes
- No
- Partially

How satisfied are you generally with the answers provided by the chat bot in terms efficiency <sup>\*</sup> and accuracy?

- 0-30%
- 31-50%
- 51-80%
- 81-100%

In which of the following categories did the chatbot perform poorly in terms of the correctness and usefulness of the answers provided? \*

- Academic Related questions
- Accommodation related questions
- Admission inquiry related questions
- History, about the institution and staff profile related questions
- Other...

Did the chatbot answer your questions promptly and correctly? \*

- Yes
- No
- Partially
- Highly dissatisfied
- Other...

Which features are most valuable to you. \*

- The Interactivity, Choice of design and Display of Output
- The Promptness of responses provided
- The accuracy of the responses provided
- Others
- None

Did the chatbot software contribute to a more efficient and time-saving inquiry session for you? \*

- Yes
- No

How well did the chatbot software adapt to your individual needs and learning style within the educational context? \*

- Perfectly
- Average
- Poorly

Overall, how likely are you to recommend this chatbot software to other students or educators in an educational setting? \*

- Certainly
- Highly likely
- Unlikely
- Other...

What important features do you think are missing in the Chatbot?

Long-answer text

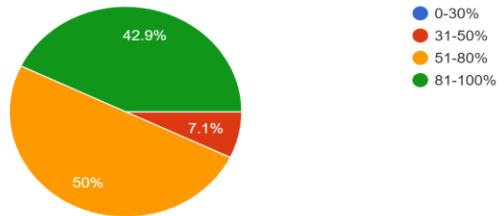
---

## Appendix III

### User Satisfaction Evaluation Questionnaire Respondents Feedbacks

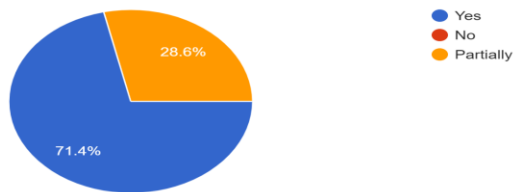
How do you rate the Vocabulary, mechanical accuracy, sentence construction and soundness of expression of the responses provided by the Chatbot?

14 responses



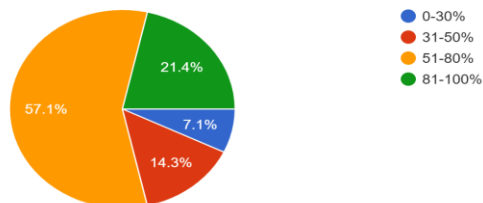
Does the system converse efficiently using your preferred language (e.g. English, French, Kindarwanda etc) ?

14 responses



How satisfied are you generally with the answers provided by the chat bot in terms efficiency and accuracy?

14 responses



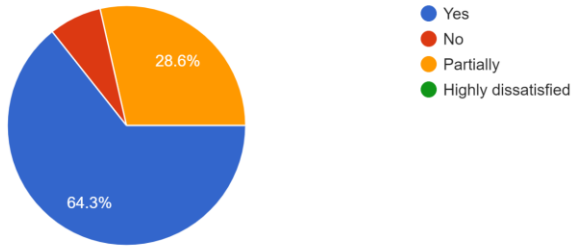
In which of the followcategories did the chatbot perform poorly in terms of the correctness and usefulness of the answers provided?

14 responses



Did the chatbot answer your questions promptly and correctly?

14 responses



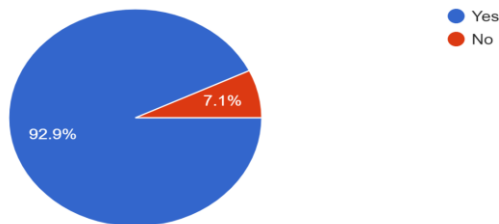
Which features are most valuable to you.

14 responses



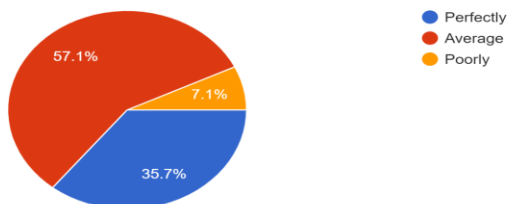
Did the chatbot software contribute to a more efficient and time-saving inquiry session for you?

14 responses

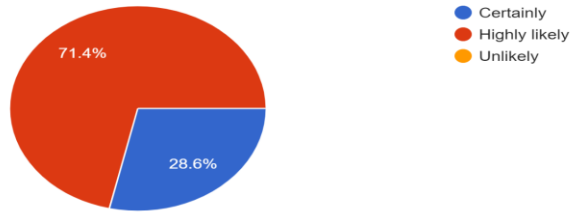


How well did the chatbot software adapt to your individual needs and learning style within the educational context?

14 responses



Overall, how likely are you to recommend this chatbot software to other students or educators in an educational setting?  
14 responses



*Figure 4.5: Users Satisfaction Experience with the research solution*

## Turnitin Plagiarism Rate

### MSC DISSERTATION

#### ORIGINALITY REPORT

**18%**

SIMILARITY INDEX

**14%**

INTERNET SOURCES

**7%**

PUBLICATIONS

**9%**

STUDENT PAPERS

#### PRIMARY SOURCES

<b>1</b>	<b>Submitted to The Robert Gordon University</b> Student Paper	<b>2%</b>
<b>2</b>	<b>ijrpr.com</b> Internet Source	<b>2%</b>
<b>3</b>	<b>github.com</b> Internet Source	<b>1%</b>
<b>4</b>	<b>www.researchgate.net</b> Internet Source	<b>1%</b>
<b>5</b>	<b>link.springer.com</b> Internet Source	<b>1%</b>
<b>6</b>	<b>Submitted to Coventry University</b> Student Paper	<b>1%</b>
<b>7</b>	<b>educationusa.state.gov</b> Internet Source	<b>&lt;1%</b>
<b>8</b>	<b>www.arxiv-vanity.com</b> Internet Source	<b>&lt;1%</b>
<b>9</b>	<b>Submitted to University of Sheffield</b> Student Paper	<b>&lt;1%</b>
<b>10</b>	<b>www.mdpi.com</b> Internet Source	<b>&lt;1%</b>

11	Submitted to University of Rwanda Student Paper	<1 %
12	www.cpur.in Internet Source	<1 %
13	arxiv.org Internet Source	<1 %
14	www.ncbi.nlm.nih.gov Internet Source	<1 %
15	Submitted to Nanyang Technological University Student Paper	<1 %
16	www.ijrcst.org Internet Source	<1 %
17	Submitted to Hills College Student Paper	<1 %

18	preview.hindawi.com Internet Source	<1 %
19	medium.com Internet Source	<1 %
20	www.ijraset.com Internet Source	<1 %
21	tudonghoa.caothang.edu.vn	

Internet Source

<1 %

22	Zeyad Ghaleb Al-Mekhlafi, Ebrahim Mohammed Senan, Taha H. Rassem, Badiea Abdulkarem Mohammed et al. "Deep Learning and Machine Learning for Early	<1 %
----	---	------

Detection of Stroke and Haemorrhage",  
Computers, Materials & Continua, 2022

Publication

---

23 Submitted to Asia Pacific University College of  
Technology and Innovation (UCTI) <1 %

Student Paper

---

24 Submitted to SP Jain School of Global  
Management <1 %

Student Paper

---

25 [huggingface.co](https://huggingface.co) <1 %

Internet Source

---

26 Submitted to Wawasan Open University <1 %

Student Paper

---

27 Submitted to Islamic University of Maldives <1 %

Student Paper

---

25 [huggingface.co](https://huggingface.co) <1 %

Internet Source

---

26 Submitted to Wawasan Open University <1 %

Student Paper

---

27 Submitted to Islamic University of Maldives <1 %

Student Paper

---

28 [www.zippia.com](https://www.zippia.com) <1 %

Internet Source

---

29 Submitted to Intercollege <1 %

Student Paper

---

30 Submitted to Tilburg University <1 %

Student Paper

30	Submitted to Humber University Student Paper	< 1 %
31	<a href="http://www.aclweb.org">www.aclweb.org</a> Internet Source	< 1 %
32	<a href="http://www.coursehero.com">www.coursehero.com</a> Internet Source	< 1 %
33	<a href="http://www.researchsquare.com">www.researchsquare.com</a> Internet Source	< 1 %
34	Dwaipayan Bandopadhyay, Rajdeep Ghosh, Rajdeep Chatterjee, Nabanita Das, Bikash Sadhukhan. "Speech Recognition and Neural Networks based Talking Health Care Bot (THCB): Medibot", 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), 2023 Publication	< 1 %

### Communication (ICCMC), 2023

Publication

35	Sruthy Anand, Moturi Karthikeya, A.M. Abhishek Sai, Ommi Balamurali. "Multi-Lingual Hybrid Chatbot for Empowering Rural Women Self-Help Groups in India", 2023 International Conference for Advancement in Technology (ICONAT), 2023 Publication	< 1 %
----	---	-------

36	Submitted to University of Surrey Student Paper	< 1 %
----	--	-------

37	<a href="http://ieeexplore.ieee.org">ieeexplore.ieee.org</a>	
----	--	--

Internet Source

< 1 %

Internet Source

<1 %

38

Partha Pratim Ray. "ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope", Internet of Things and Cyber-Physical Systems, 2023

Publication

<1 %

39

fdocuments.net

Internet Source

<1 %

40

medicmind.com.au

Internet Source

<1 %

41

Submitted to Vels University

Student Paper

<1 %

42

dokumen.pub

Internet Source

<1 %

43

medinform.jmir.org

Internet Source

<1 %

44

www.hindawi.com

Internet Source

<1 %

45

"AI, Blockchain and Self-Sovereign Identity in Higher Education", Springer Science and Business Media LLC, 2023

Publication

<1 %

46

Submitted to Koc University

Student Paper

47 Submitted to National Institute of Technology, Raipur <1 %  
Student Paper

---

48 [www.yumpu.com](http://www.yumpu.com) <1 %  
Internet Source

---

49 Submitted to aun <1 %  
Student Paper

---

50 [digitalcommons.unl.edu](http://digitalcommons.unl.edu) <1 %  
Internet Source

---

51 [ediss.sub.uni-hamburg.de](http://ediss.sub.uni-hamburg.de) <1 %  
Internet Source

---

52 [mhealth.jmir.org](http://mhealth.jmir.org) <1 %  
Internet Source

---

53 Submitted to De Montfort University <1 %  
Student Paper

---

54 Submitted to Sreenidhi International School <1 %  
Student Paper

---

55 [eprint.iitd.ac.in](http://eprint.iitd.ac.in) <1 %  
Internet Source

---

56 [isset.shiksha.com](http://isset.shiksha.com) <1 %  
Internet Source

---

57 Submitted to Higher Education Commission Pakistan <1 %  
Student Paper

---

58 [hr.mcleanco.com](http://hr.mcleanco.com) <1 %  
Internet Source

---

59 repository.smuc.edu.et Internet Source <1 %

60 www.unlv.edu Internet Source <1 %

61 Elizabeth C. Reilly. "Rwandan Women in Higher Education: Progress, Successes, and Challenges", Frontiers in Education, 2021 Publication <1 %

62 busman.mandela.ac.za Internet Source <1 %

63 www.massagetherapynearme.com Internet Source <1 %

64 Mohammad Shahin, F. Frank Chen, Ali Hosseinzadeh, Neda Zand. "Using Machine Learning and Deep Learning Algorithms for Downtime Minimization in Manufacturing Downtime Minimization in Manufacturing Systems: An Early Failure Detection Diagnostic Service", Research Square Platform LLC, 2023 Publication <1 %

iceb.johogo.com

65 Internet Source <1 %

66 scholarshare.temple.edu Internet Source <1 %

67 www.sitel.com Internet Source <1 %

68 Markus Orsi, Jean-Louis Reymond. "GPT-3 accurately predicts antimicrobial peptide <1 %

69	Submitted to University of Wales Swansea Student Paper	<1 %
70	marielebert.wordpress.com Internet Source	<1 %
71	tudr.thapar.edu:8080 Internet Source	<1 %
72	ulspace.ul.ac.za Internet Source	<1 %
73	www.bestvalueschools.org Internet Source	<1 %
74	www.futureknowledgework.com Internet Source	<1 %
75	Abid Haleem, Mohd Javaid, Ravi Pratap Singh. "An era of ChatGPT as a significant futuristic	<1 %

support tool: A study on features, abilities,  
and challenges", BenchCouncil Transactions  
on Benchmarks, Standards and Evaluations,  
2023

Publication

76	Submitted to Liberty University Student Paper	<1 %
77	Safari Mukeru. "On the Convergence of Series of Dependent Random Variables", Journal of Theoretical Probability, 2020 Publication	<1 %
78	Submitted to University of Leeds Student Paper	<1 %
79	v-scheiner.brunel.ac.uk Internet Source	<1 %

80	vital.seals.ac.za:8080 Internet Source	<1 %
81	www.science.gov Internet Source	<1 %
82	www.siddkaramcheti.com Internet Source	<1 %
83	Dean Fantazzini. "Default forecasting for small-medium enterprises: does heterogeneity matter?", International Journal of Risk Assessment and Management, 2009 Publication	<1 %
	Shot Learning", Cold Spring Harbor Laboratory, 2023 Publication	
85	Hans W. Ittmann. "The current state of OR in Africa", Operational Research, 2019 Publication	<1 %
86	J. M. Mesa, C. Menendez, F. A. Ortega, P. J. Garcia. "A smart modelling for the casting temperature prediction in an electric arc furnace", International Journal of Computer Mathematics, 2009 Publication	<1 %
87	Submitted to Liverpool John Moores University Student Paper	<1 %
88	Submitted to University College London Student Paper	<1 %

---

89 [arrow.tudublin.ie](http://arrow.tudublin.ie) <1 %  
Internet Source

---

90 [ebin.pub](http://ebin.pub) <1 %  
Internet Source

---

91 [essay.utwente.nl](http://essay.utwente.nl) <1 %  
Internet Source

---

---

92 [itnt-conf.org](http://itnt-conf.org) <1 %  
Internet Source

---

93 [mafiadoc.com](http://mafiadoc.com) <1 %  
Internet Source

---

94 [repository.tudelft.nl](http://repository.tudelft.nl) <1 %  
Internet Source

---

95 [tel.archives-ouvertes.fr](http://tel.archives-ouvertes.fr) <1 %  
Internet Source

---

96 [whiceb.cug.edu.cn](http://whiceb.cug.edu.cn) <1 %  
Internet Source

---

97 [www.sbir.gov](http://www.sbir.gov) <1 %  
Internet Source

---

98 [www.scilit.net](http://www.scilit.net) <1 %  
Internet Source

---

99 S. Aslani, J. Jacob. "Utilisation of deep learning for COVID-19 diagnosis", Clinical Radiology, 2023  
Publication

---

100 "Advances in Computing and Data Sciences", Springer Science and Business Media LLC, 2022 <1 %

---