OMAR-HUSSAIN COMPARATIVE ANALYSIS OF ARTIFICIAL INTELLIGENCE CHATBOTS MASTER THESIS 2023

NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF COMPUTER INFORMATION SYSTEMS

COMPARATIVE ANALYSIS OF ARTIFICIAL INTELLIGENCE CHATBOTS

M.Sc. THESIS

Omar HUSSAIN

Nicosia

June 2023

NEAR EAST UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF COMPUTER INFORMATION SYSTEMS

COMPARATIVE ANALYSIS OF ARTIFICIAL INTELLIGENCE CHATBOTS

M.Sc. THESIS

Omar HUSSAIN

Supervisor

Professor Dr Nadire CAVUS

Nicosia

June 2023

Approval

We certify that we have read the thesis submitted by Omar Hussain titled **"Comparative Analysis of Artificial Intelligence Chatbots"** and that in our combined opinion, it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Educational Sciences.

Examining Committee	**Name-Surname**	**Signature**
Head of the Committee:	Prof. Dr. Fezile Özdamlı	.....
Committee Member*:	Assist. Prof. Dr. Sahar Ebadinezhad	.....
Supervisor:	Prof. Dr. Nadire Cavus	.....

Approved by the Head of the Department

05/07/2023

.....

Prof. Dr. Nadire Cavus

Head of Department

Approved by the Institute of Graduate Studies

.../2023

.....

Prof. Dr. Kemal Hüsnü Can Başer

Head of the Institute

Declaration

I hereby declare that all information, documents, analysis, and results in this thesis have been collected and presented according to the academic rules and ethical guidelines of Institute of Graduate Studies, Near East University. I also declare that as required by these rules and conduct, I have fully cited and referenced information and data that are not original to this study.

OMAR HUSSAIN

…../…../…..

Acknowledgments

I would like to express my sincere gratitude and appreciation to Professor Doctor Nadire Cavus for her invaluable guidance, support, and mentorship throughout the course of this research and my entire study path at Near East University. Her expertise, insightful feedback, and unwavering encouragement have been instrumental in shaping this study. I am truly fortunate to have had the opportunity to work under her supervision and benefit from her vast knowledge in the field. And I will give thanks to almighty for keeping me healthy and strong throughout my study period. Lastly, I am grateful to my family, friends, and colleagues for their understanding, patience, and continuous encouragement during this research journey. Their unwavering support has been a constant source of motivation.

Omar Hussain

Abstract

Comparative Analysis of Artificial intelligence Chatbots

Hussain, Omar

MA, Department of Computer Information Systems

July 2023, 81 pages

The surge in the adoption of chatbot technologies is attributed to their significant capacity to enhance customer service, support, and engagement in various fields such as healthcare, education, and e-commerce. However, the success of chatbots is dependent upon their ability to effectively comprehend natural language, acquire knowledge from user interactions, and provide engaging user experiences. The aforementioned factors are significant in determining the overall usefulness and functionality of chatbots in addressing the evolving requirements of their users. This research was conducted as a comparative analysis of the performance of several chatbots, that include Chat GPT, MS Bing Chat, ChatSonic AI, Character AI, YouChat AI, Perplexity AI, and Google Bard. Examining factors involving accuracy, response speed, conversation flow, adaptability, personalization, user satisfaction, robustness, security, scalability, and integration. Results showed Chat GPT excelled in accuracy and conversation flow, while MS Bing Chat showed exceptional response speed. ChatSonic AI displayed high adaptability and personalization, while YouChat AI received positive feedback. While Google Bard demonstrated strong scalability and integration. Improving robustness, security, adaptability, and personalization is recommended. The study emphasizes considering multiple metrics and suggests research avenues like advanced natural language processing techniques and robust security frameworks.

Key Words: chatbots, natural language understanding, natural language processing, artificial intelligence chatbots, user experience.

Özet

Comparative analysis of Artificial Intelligence Chatbots

Hussain, Omar

MA, Department of Computer Information Systems

July 2023, 81 pages

Chatbot teknolojilerinin benimsenmesinin artmasının nedeni, sağlık hizmetleri, eğitim ve e-ticaret gibi çeşitli sektörlerde müşteri hizmetleri, destek ve etkileşimi artırma kapasitelerine sahip olmalarıdır. Chatbotların başarısı, doğal dilin iyi anlaşılması, kullanıcı etkileşimlerinden bilgi edinme yeteneği ve etkileyici kullanıcı deneyimleri yaratma yeteneğine bağlıdır. Chatbotların kullanıcıların değişen ihtiyaçlarına nasıl yanıt vereceğini belirlemek için bu kriterler önemlidir. Bu çalışmanın amacı, Chat GPT, MS Bing Chat, ChatSonic AI, Character AI, YouChat AI, Perplexity AI ve Google Bard gibi çeşitli chatbotların performanslarını karşılaştırmaktır. Ölçeklenebilirlik, entegrasyon, doğruluk, yanıt hızı, konuşma akışı, uyarlanabilirlik, kullanıcı memnuniyeti, sağlamlık, güvenlik ve ölçeklenebilirlik dahil olmak üzere çeşitli bileşenler incelendi. Sonuçlar, Chat GPT'nin doğruluk ve konuşma akışında mükemmel olduğunu ve MS Bing Chat'in olağanüstü yanıt hızını gösterdi. YouChat AI olumlu geri bildirimler aldı, ancak ChatSonic AI daha kişiselleştirilmiş ve uyarlanabilirdi. Google Bard, aynı zamanda güçlü ölçeklenebilirlik ve entegrasyon sergiledi. Sağlamlık, güvenlik, uyarlanabilirlik ve özelleştirmenin geliştirilmesi önerilir. Araştırma, gelişmiş doğal dil işleme teknikleri ve sağlam güvenlik çerçeveleri gibi araştırma alanlarının araştırılması gerektiğini vurgulamaktadır.

***Anahtar Kelimeler**:* chatbotlar, doğal dil anlama, doğal dil işleme, yapay zeka chatbotları, kullanıcı deneyimi.

Table of Contents

CHAPTER III

CHAPTER IV

CHAPTER V

CHAPTER VI

List of Tables

List of Figures

List of Abbreviations

A.I: Artificial Intelligence

NLP: Natural Language Processing

ML: Machine Learning

RNN: Recurrent Neural Networks

GPT: Generative Pre-trained Transformer

UX: User Experience

MIT: Massachusetts Institute of Technology

DM: Dialogue management

UI: User Interface

MCDA: Multi-Criteria Decision analysis

CNN: Convolutional Neural Network

BLEU: Bilingual Evaluation Understudy

F1 score: F-score or F-measure is a measure of a test's accuracy

IBM: The International Business Machines

LaMDA: Language Model for Dialogue Applications

CHAPTER I
INTRODUCTION

This chapter functions as an exhaustive overview of the entire research, with an emphasis on the problem statement, aims of the study, and contributions of the research. In addition, it describes the distinctive aspects of the research and offers solutions to specific problems identified during the investigation. By delving into these facets, this chapter provides a comprehensive overview of the research endeavour, emphasising its significance and highlighting the steps taken to resolve any obstacles encountered.

1.1 Background

Artificial Intelligence (A.I) chatbots have emerged as a significant technological advancement in recent years, revolutionizing the way businesses and individuals interact online (Kumar, 2021). These intelligent conversational agents (Gao & Jiang, 2021; Nicolescu & Tudorache, 2022) are designed to simulate human-like conversations and provide automated responses, delivering personalized assistance and support to users (Camilleri & Troise, 2023). The widespread adoption of chatbot technology, powered by A.I and natural language processing (NLP) (Olujimi & Ade-Ibijola, 2023), has transformed various industries, including customer service (Cordero et al., 2022; Haugeland et al., 2022), education and research (Kooli, 2023), finance, entertainment, healthcare, and information retrieval (Caldarini et al., 2022; Vijayaraghavan et al., 2020). A.I chatbots have emerged as valuable tools for enhancing user interactions, automating tasks, and improving overall user experiences (Xu et al., 2021a).

Early work in NLP and Machine Learning (ML) laid the groundwork for today's A.I chatbots. NLP is the study of how computers can read, process, and even create human language. Using ML methods, computers may enhance their performance by learning from data on their own (Adamopoulou & Moussiades, 2020). The first generation of chatbots were rule-based, meaning that their replies to user inputs were predetermined by a set of rules (Caldarini et al., 2022). They were useful for the most part, but their answers were simplistic, and they couldn't process complicated questions. Chatbots, however, have advanced greatly because of the developments in A.I and ML (Adam et al., 2021a).

The capabilities of today's AI chatbots have been significantly improved by combining NLP, ML algorithms, and deep learning strategies (Lin et al., 2023). They make use of models such as Recurrent Neural Networks (RNNs), Transformers, and Generative Pre-trained Transformer (GPT) models, all of which have significantly contributed to improvements in language interpretation and creation (Han et al., 2021; Wang et al., 2022). These chatbots, which are driven by A.I, now have the ability to grasp the intent, context, and sentiment of the user, which enables them to offer replies that are more accurate and pertinent. They are able to respond to a broad variety of questions, offer assistance with activities, provide individualised advice, and even carry-on discussions that are both natural and meaningful (Haleem et al., 2022a).

A.I chatbots are becoming increasingly popular among businesses in a wide variety of sectors as helpful tools for customer care, sales, technical assistance, and marketing (Adam et al., 2021b). Chatbots may be incorporated into a variety of digital mediums, including websites, chat platforms, mobile applications, and voice assistants (Enholm et al., 2022). This enables organisations to offer support around the clock and to expedite their business processes. In addition, artificial intelligence chatbots continue to make progress thanks to continuing research and development initiatives (Okonkwo & Ade-Ibijola, 2021). They are continually enhancing their performance by gaining knowledge through interactions with users and drawing insights from data in order to become more effective, accurate, and capable of comprehending the requirements and preferences of users. In order to teach chatbots to interact in a more natural way and to optimise their replies depending on the input they get from users, reinforcement learning techniques are being utilised (Andrade & Tumelero, 2022).

A.I chatbots have provided countless benefits, but there are still many obstacles to overcome. A very important factor to take into account is programming chatbots to behave in a way that is both ethical and objective (Kooli, 2023). Concerns have been expressed regarding the possibility for the transmission of social prejudices and disinformation due to the existence of bias in training data as well as improper responses (King & chatGPT, 2023). Researchers are currently working on establishing approaches to increase the transparency and accountability of chatbot systems, as well as to reduce the impact of bias on such systems. Keeping private information safe while being

handled is still another difficulty (Hasal et al., 2021). Because chatbots frequently deal with sensitive information, extensive security measures are required to safeguard the privacy of users (Sebastian, 2023). A safe chatbot deployment must always include crucial components like as encryption, authentication, and compliance with any data protection standards.

Achieving an optimal equilibrium between automation and human intervention is a crucial factor to take into account (Misischia et al., 2022). Although chatbots are capable of autonomously addressing a diverse array of inquiries, there exist situations where human intervention becomes imperative. The integration of live chat, which facilitates the seamless transition from chatbots to human agents, is a crucial aspect of customer service (Adam et al., 2021a). This approach guarantees a seamless customer experience and effectively addresses intricate issues that necessitate human expertise.

1.2 Problem Statement

One of the primary challenges in chatbot development lies in the evaluation and comparison of different chatbot systems to properly choose the right one for use in areas where they're needed or improve their capabilities in respect to the developing trends in their respective field. Despite a rise in chatbot-related literature, including an overall increase in articles employing assessment approaches, no widely accepted benchmark for chatbot evaluation has yet arisen. Existing evaluation methodologies often lack comprehensive criteria and fail to capture the multifaceted nature of chatbot performance according to Cesas et al. (2023). The study conducted by Bhardwaz & Kumar (2023) employed a mixed-methods approach, utilising both quantitative analysis of chatbot performance metrics and qualitative analysis of user feedback. However, the study was limited in scope, as it only compared three chatbots and assessed their performance based on a limited number of queries. Therefore, it is essential to implement a broad comparative approach that incorporates a broader spectrum of variables and integrates easily accessible chatbot platforms. This method has the potential to enhance and augment the scope of research within the given field. By adopting an expanded outlook and employing more user-friendly chatbot platforms, it is possible to gain fresh insights and further progress in this field, thereby expanding the frontiers of knowledge in A.I chatbot literature.

1.3 Aim of the Study

The aim of this study is to perform a comparative analysis of several chatbot systems, namely Chat GPT, MS Bing Chat, ChatSonic AI, Character AI, YouChat AI, Perplexity AI, and Google Bard. This research endeavour seeks to evaluate and compare the selected chatbots utilising an extensive range of evaluative measures, such as accuracy, response speed, conversation flow, adaptability, personalization, user satisfaction, robustness and security, scalability, and integration. Also, this study will offer valuable insights into the strengths and weaknesses of all chatbot systems by comparing and evaluating these platforms. This will facilitate informed decision-making in the selection, development, improvement, and implementation of A.I chatbots for various use cases.

1.4 Contribution

This study contributes to the advancement of chatbot evaluation and development through various means. Firstly, this study will contribute to the areas of A.I chatbots research in the department of computer information systems (CIS) of the university by providing insightful knowledge to future researchers based on the findings. Fundamentally, it provides a system for comparison and evaluating the performance of chatbots from multiple dimensions. This paradigm effectively addresses the limitations of previous evaluation methodologies and presents a broader perspective on the capabilities of chatbots. The study presents a comparative analysis of multiple chatbot systems, elucidating significant distinctions and similarities among them. The findings of this study can serve as a guide for individuals who may need them for personal use, researchers who are keen to dive into the A.I chatbot research, developers, corporations, and organisations seeking to establish a successful chatbot selection, usage, improvement, and development strategy. This study makes a valuable contribution to the academic field of artificial intelligence and natural language processing by enhancing the understanding of chatbot technology and its potential applications.

1.5 Overview of the Thesis

This thesis is broken down into five major segments for easier reading. The first chapter introduces the topic of the research, presents the problem statement, and defines the purpose of the study as well as the contributions that it makes to all areas of interest.

The second chapter does a thorough literature review on topics pertaining to chatbots technology, assessment methodology, and existent chatbot systems.

The methodology that was utilised in this study is outlined in Chapter 3, which covers topics such as the selection of chatbot systems, implementation processes, data collection processes, and assessment criteria to be used on all selected chatbots used in this study.

In Chapter 4, the in-depth results and comparative analysis of all the chatbots that were chosen based on the predetermined criteria are presented.

The findings, their consequences, and some recommendations for further research are presented and discussed in Chapter 5.

Finally, the research is concluded in chapter 6, which also provides recommendations for further study and to end users.

CHAPTER II

LITERATURE REVIEW

The literature review chapter of this study offers an in-depth review of prior research and academic literature relevant to the field of comparative analysis of A.I chatbots. The objective of this review is to analyse the present state of knowledge regarding the assessment of chatbot performance, user satisfaction, and the diverse aspects of chatbot functionality. Through the process of synthesising and analysing a diverse array of literature, this section offers valuable insights into the strengths, weakness, and gaps in the existing literature. Furthermore, it acts as a basis for comprehending the theoretical and conceptual frameworks that steer this comparative analysis. The review of literature serves the purpose of not only setting the stage for the research but also serves as a foundation for developing research inquiries, recognising research gaps, and moulding the research approach.

2.1 Theoretical Framework

2.1.1 Artificial Intelligence

Artificial intelligence (A.I) is a discipline within the field of computer science that is concerned with the development of intelligent machines capable of carrying out tasks that would typically necessitate human intelligence, such as visual perception, speech recognition, decision-making, and language translation. The field of A.I encompasses a diverse array of technologies and applications (Mukhamediev et al., 2022). Fundamentally, A.I pertains to the development of computational systems that are capable of executing functions that conventionally necessitate human cognitive abilities, such as acquiring knowledge, resolving complex issues, and making informed judgements (Xu et al., 2021b). The significance of A.I has grown substantially across various domains such as healthcare, transportation, finance, and environmental monitoring, as noted by (Perifanis & Kitsios, 2023).

The origins of A.I can be traced to the 1950s, when scholars at academic institutions and research facilities began to acknowledge the capacity of computers to emulate human intelligence (Chellappa, 2022). Alan Turing, in 1950, played a significant role in the establishment of contemporary computers and artificial intelligence. The

Turing test was predicated upon the notion that a computer's intelligent conduct is measured by its capacity to attain human-like proficiency in cognitive tasks (Amisha et al., 2019). The term "Artificial Intelligence" was introduced by John McCarthy, a renowned researcher in the field of A.I, during a conference that took place at Dartmouth College in New Hampshire in 1956 (Cordeschi, 2007). During the initial stages, researchers in the field of artificial intelligence were predominantly concentrated on the creation of rule-based systems and knowledge bases with the aim of resolving intricate issues (Xu et al., 2021b). The domain of A.I encountered a significant impediment during the 1970s as scholars acknowledged that the creation of all-purpose A.I systems turned out to be considerably more difficult than they had initially envisioned. This particular era is commonly referred to as the "A.I winter" within academic circles (Huawei Technologies Co., Ltd., 2023). During the period spanning the 1990s and early 2000s, the development of artificial intelligence was primarily propelled by the progression of machine learning and data mining approaches. The aforementioned technologies have facilitated the ability of A.I systems to acquire knowledge from vast quantities of data and generate prognostications through identification of patterns within said data (Sarker, 2022). A.I applications have demonstrated tangible achievements in practical domains, including but not limited to speech recognition, computer vision, and NLP (Guan et al., 2020).

For a considerable period, the domain of A.I was primarily centred on theoretical aspects, with limited practical implications (Xu et al., 2021b). Over the last decade, significant changes have occurred in the field of ML due to the convergence of several factors, including the increased processing power of machines, enhanced learning algorithms, and improved accessibility to large datasets (Sarker, 2021). These developments have facilitated the advancement of ML and its widespread adoption in various industries (Linardatos et al., 2020). A.I possesses the capability to revolutionise various domains of human existence, such as transportation, energy, agriculture, and education (Sarker, 2021). Notwithstanding, ethical considerations regarding A.I-driven systems, including but not limited to biases, privacy, and security, have also been raised (Naik et al., 2022). As the field of A.I advances, it becomes increasingly crucial to ensure

that AI systems are developed and implemented in manners that prioritise both safety and ethical considerations (Pflanzer et al., 2022).

The advancement of machine learning algorithms has been a significant catalyst for the progression of artificial intelligence, as it allows computers to acquire knowledge from data and utilise it to make informed decisions by identifying patterns and trends. The field of machine learning encompasses a variety of algorithmic approaches, among which are supervised learning, unsupervised learning, and reinforcement learning (Sarker, 2021; Zhou et al., 2023). The process of supervised learning entails the training of a machine learning model on data that has been labelled, whereas unsupervised learning involves the identification of patterns and structures within data that has not been labelled (Pugliese et al., 2021). Reinforcement learning is a machine learning paradigm that entails the process of instructing a model to make decisions by leveraging feedback obtained from the environment (Habehh & Gohel, 2021; Zhao et al., 2023).

NLP is a significant field of AI research that encompasses the creation of algorithms and technologies capable of comprehending and producing natural language. NLP encompasses a diverse range of applications, such as chatbots, speech recognition, and language translation (Khurana et al., 2023). A primary obstacle encountered in NLP pertains to managing the intrinsic ambiguity and intricacy of language, alongside the variances in cultural and linguistic aspects across diverse languages (Davenport & Kalakota, 2019).

A.I is an essential component in the development of chatbots, specifically in the areas of NLP and user experience (UX). NLP pertains to the capacity of a machine to comprehend and decipher human language, a crucial aspect for chatbots to establish effective communication with users (Adamopoulou & Moussiades, 2020). Chatbots that are powered by A.I have the capability to utilise NLP techniques to scrutinise user input and furnish precise and contextually relevant responses, thereby enhancing the overall user experience (Caldarini et al., 2022). Furthermore, A.I has the potential to customise chatbot interactions by taking into account user preferences and behaviour, thereby augmenting the overall user experience according to Zhang et al. (2020). The assessment of conversational A.I systems' efficacy can be approached from four distinct viewpoints: user experience, information retrieval, linguistic, and artificial intelligence (Jadeja &

Varia, 2017). A.I has the potential to enhance chatbot performance by enabling them to learn from user interactions and subsequently improve their effectiveness and efficiency is an important point highlighted by Kooli (2023). It is crucial to guarantee the responsible and ethical development and utilisation of A.I-driven chatbots, while considering factors such as privacy, bias, and job displacement as reported by Xu et al. (2021b). Although A.I presents numerous potential advantages, it is crucial to consider significant ethical and societal implications (Bankins & Formosa, 2023). A primary area of concern pertains to the possible implications on employment, given the capacity of A.I and automation technologies to supplant numerous jobs and industries. A.I technologies have raised concerns regarding privacy, security, and bias due to their ability to gather and scrutinise vast amounts of data, frequently without individuals' awareness or authorization (Davenport & Kalakota, 2019).

2.1.2 Chatbots

Chatbots are software applications that are specifically created to mimic human dialogue by means of text or voice-based interactions. NLP and ML algorithms are utilised for comprehending and addressing user queries as defined by Adamopoulou & Moussiades (2020). Chatbots possess the capability to be incorporated into a multitude of platforms such as websites, messaging applications, and social media platforms. This integration enables them to furnish customer support, automate tasks, and enhance user engagement (Adamopoulou & Moussiades, 2020).

The origin of chatbots can be traced back to the 1960s, when Joseph Weizenbaum of Massachusetts institute of technology (MIT) developed the first chatbot, ELIZA (Casas et al., 2020). ELIZA was a chatbot that employed a rule-based approach to engage in a simulated dialogue with a human, utilising pattern matching techniques. The system was developed with the intention of emulating the role of a psychotherapist, and had the capability to generate queries and offer compassionate feedback in response to user interactions (Natale, 2019). During the 1970s and 1980s, the evolution of chatbots progressed as a result of the advancement of more advanced NLP methods and the implementation of expert systems. Expert systems were computer programmes that operated on the basis of rules and utilised knowledge bases to offer guidance and

suggestions to users (Caldarini et al., 2022). In the 1990s, firms began utilising chatbots to automate customer service and minimise response times. Early chatbots couldn't grasp real language, which frustrated users (Haugeland et al., 2022). In the 2000s and 2010s, ML techniques enabled chatbots to learn from user interactions in order to enhance their feedback. NLP and ML algorithms let ML-based chatbots interpret and answer user inquiries (Caldarini et al., 2022).

The two primary classifications of chatbots are rule-based and A.I-powered. Chatbots that are rule-based function based on a predetermined set of rules and are only capable of responding to a set of commands or instructions (Adamopoulou & Moussiades, 2020a). Conversational agents that are powered by A.I utilise techniques such as ML and NLP to comprehend and provide answers to a diverse array of user queries (Mariani et al., 2023).

The rapid and effective provision of customer support and service is considered a significant benefit of chatbots. Chatbots have the potential to enhance customer satisfaction by streamlining response times through the automation of routine tasks and interactions (Haleem et al., 2022a). Chatbots have the potential to gather valuable data and insights pertaining to user behaviour and preferences, thereby facilitating the enhancement of products and services (Xu et al., 2021b). Chatbots have the capability to be integrated with various systems and applications, thereby enabling the automation of intricate tasks and facilitating the optimisation of workflows (Adamopoulou & Moussiades, 2020a).

Notwithstanding their potential benefits, chatbots are also subject to certain limitations and concerns (Haugeland et al., 2022). One of the primary obstacles is to ensure that chatbots possess the capability to comprehend and react to users in a manner that is both natural and intuitive, given that users frequently exhibit distinct communication styles and preferences (Li et al., 2021). Furthermore, chatbots are prone to inaccuracies and misinterpretations, and they may lack the capability to manage unforeseen or intricate circumstances (Parviainen & Rantala, 2022). Regardless those hurdles, it is probable that chatbots will persist in fulfilling a significant function across various sectors and implementations in the forthcoming years (Haleem et al., 2022b). With the advancement of technology, chatbots are expected to enhance their capacity to

comprehend and address user queries, thereby augmenting their significance as a customer service and support mechanism.

2.1.3 Artificial Intelligence Chatbots

AI chatbots are software programmes that utilise A.I and NLP techniques to simulate human-like conversation through interfaces that operate on either text-based or voice communication (Adamopoulou & Moussiades, 2020b). A.I systems possess the ability to comprehend and react to human language input, rendering them a progressively favoured resource for corporations and institutions seeking to optimise their customer service endeavours (Dwivedi et al., 2022). Chatbots can be enabled by NLP algorithms, which facilitate their comprehension and reaction to human language input (Adamopoulou & Moussiades, 2020a). The integration of other AI technologies, such as machine learning, can enhance the responsiveness of chatbots to user input (Okonkwo & Ade-Ibijola, 2021). Chatbots are offered in diverse formats, such as voice-activated assistants, messaging applications, and website plugins, and can be coded to execute a broad spectrum of functions, ranging from responding to basic inquiries to managing intricate customer service demands.

A significant hurdle in the development of AI chatbots pertains to the assurance of precision and dependability of the information dispensed by these chatbots, given that their efficacy is contingent upon the quality of the data they are trained on (Loh, 2023). An additional obstacle involves guaranteeing the confidentiality and protection of delicate data. Moreover, there exist apprehensions regarding conversational partiality arising from detrimental training data. These concerns can be resolved by designing filtration algorithms that evaluate the toxicity degree of a chatbot's reply and exclude statements that exceed a pre-established threshold of conversational partiality (Alhajjar & Bradley, 2022).

Despite the considerable advancements in chatbot technology, their proficiency in handling nuanced and contextual customer service interactions still falls short of that of human representatives. Consequently, it is possible that chatbots may not be appropriate for certain customer service contexts, particularly those that necessitate a significant degree of individualization or emotional intelligence (Caldarini et al., 2022).

2.2 Related Research

The author of this comparative study expanded upon the results of several prior research relating to the subject matter, with the objective of evaluating the performance of a variety of A.I chatbots. The comprehension and techniques have been considerably influenced by the significant sources emphasized in this segment.

By combining quantitative analysis of chatbot performance metrics with qualitative analysis of user feedback, Bhardwaz & Kumar (2023) conducted a similar study to compare the performance of three chatbot technologies in the areas of NLP, ML techniques, and user experience. Accuracy, reaction speed, relevance, user satisfaction, and user engagement were some of the factors studied. The study found considerable variations in the effectiveness criteria used to evaluate the various chatbot systems. Results also indicated that Google BARD had the quickest reaction time among the chatbot technologies tested, while ChatGPT ranked highest in terms of accuracy and relevancy. The highest levels of user satisfaction and activity were seen for Microsoft Bing. The research also highlighted the value of NLP and ML for improving chatbot efficiency. The study's results have significant bearing on the design and use of chatbots. Each chatbot technology's advantages and disadvantages were studied, shedding light on how each may be enhanced. The authors identify the creation of more sophisticated NLP algorithms as a potentially fruitful area for future study. The authors also noted that existing algorithms have made great gains towards enhancing chatbot performance, but that further development is still necessary. The authors went on to point out that chatbots, in particular, may use improvements in their capacity to understand and reply to inquiries that have several parts. Further study might also investigate methods for enhancing the precision of NLP in industry-specific chatbots, such as those used in healthcare and finance.

Peyton & Unnikrishnan (2023) specifically focused their research in comparison on the effectiveness of popular chatbot platforms and a state-of-the-art Sentence BERT (SBERT) implementation in the context of educational domain-specific datasets. The research evaluates the state-of-the-art Sentence BERT (SBERT) model used to construct a reliable chatbot in comparison to the outcomes of intent classification using two widely used chatbot frameworks Feedforward model, and an SBERT model. The researchers'

findings revealed the comparative performance of these chatbot technologies, shedding light on their potential applications in addressing the needs of prospective students within online learning environments. Such studies contribute to a deeper understanding of the current landscape of chatbot research and highlight the advances being made in this rapidly evolving field.

According to the comparative analysis carried out by Sharma et al. (2020), the authors compare the features of five widely used chatbot platforms Dialogflow, Amazon Lex, the international business machines (IBM) Watson Assistant, Microsoft Bot Framework, and Rasa with respect to NLP Dialogue management(DM) (Suta et al., 2020), and user interface(UI). The article examines the platforms based on a number of criteria, including accuracy, usability, scalability, and cost. The authors also give advice on how to select the best platform for the chatbot software, taking into account its unique needs. The research found that while each platform has its benefits and drawbacks, Dialogflow and Rasa are the best for creating sophisticated chatbots that can respond appropriately to a broad variety of user inputs. Researchers and developers of chatbots for a wide range of uses and industries may find this article a valuable resource.

As per the comparative research by Abdellatif et al. (2022) which focused on a comparison of several systems for NLP used in chatbot development in software engineering. The authors examine four well-known NLU platforms; Dialogflow, IBM Watson, Luis.ai, and Wit.ai and compare them in terms of their precision in understanding software engineering inquiries, their user-friendliness, and their affordability. The researchers revealed that while each platform has its advantages and disadvantages, Dialogflow and Luis.ai stand out as the most effective and user-friendly options for software engineers developing chatbots. Using NLU platforms for software engineering chatbots has a number of obstacles, as the authors note, including the necessity for specific domain expertise and the limits of existing NLU models in comprehending technical terminology and jargon. While highlighting the need for more research to tackle the obstacles associated with employing NLU platforms for technical domains, the study serves as a helpful resource for scholars and practitioners interested in creating chatbots for software engineering applications.

Chakrabortty et al. (2023) suggested a unique method for assessing the differences between potential chatbot options, taking into account user fulfilment, accuracy, efficiency, and cost by combining fuzzy logic, grey theory, and multi-criteria decision analysis (MCDA). The suggested model accounts for the indeterminacy of individual users' choices and the wide range of chatbots' capabilities. An actual case study of choosing a chatbot for a customer support system is used to test the validity of the presented methodology. The outcomes demonstrate the effectiveness of the suggested strategy in dealing with ambiguity and providing a trustworthy resource for decision-makers to choose the best chatbot for their needs. For anyone working on decision-making models for choosing chatbots and other A.I-based systems in the face of uncertainty, this paper is a valuable resource. The suggested strategy can aid businesses in picking the most suitable chatbot option to enhance their customer support systems and boost their overall performance.

Accordingly, the research by Balas & Ing (2023) which evaluates the diagnostic accuracy of two conversational AI models for ocular disorders, namely ChatGPT and the Isabel Pro Differential Diagnosis Generator. The two models are compared in terms of how well they can detect and rate probable diagnoses, how well they can ask pertinent follow-up questions, and how well they can create an overall usable and enjoyable experience for the user. They test the two algorithms against a human expert using a dataset of 25 ophthalmic cases. Both ChatGPT and the Isabel Pro Differential Diagnosis Generator are found to be useful in making correct diagnosis of ocular disorders, according to the study. In contrast to Isabel Pro's Differential Diagnosis Generator, ChatGPT excels in usability and user experience by providing more pertinent follow-up questions. The authors believe that ChatGPT and similar conversational A.I models have the potential to significantly improve the quality and efficiency with which ophthalmologists get diagnostic help. This study is a valuable resource for scientists and healthcare professionals working on conversational AI models for medical diagnosis and related fields.

The research by Anki et al. (2021) made use of comparative analysis of two distinct multimodal chatbot implementations that are based on news classification data, the two implementations are compared in terms of their accuracy in categorizing news

articles into different groups and their efficacy in generating appropriate responses to user queries; the FastText algorithm is compared with a convolutional neural network (CNN). Results show that both CNN and FastText models achieve high accuracy in categorizing news stories, with CNN showing a slight edge. As a result of its superior performance in producing relevant replies to user inquiries, the FastText model is more ideally suited for usage in chatbots. The authors suggest that chatbots' accuracy and efficacy in processing complicated data may be improved by using multimodal techniques, and that the choice of algorithm and implementation methodology is crucial in designing effective chatbots for specific applications. This article is a valuable resource for academics and professionals building chatbots for use in news categorization and other NLP-based tasks.

In the comparative study of cloud platforms to develop A.I Chatbots by Patil et al. (2017), the authors comparison included cloud-based chatbot technologies with various constraints, such as built-in artificial intelligence, the amount of time required to set up, and the total time it takes to finish. In conclusion, the comparison will allow researchers to determine which of the cloud platforms is the most effective and well-suited for the development of chatbots. Their study showed that different chatbot cloud platforms had different benefits and drawbacks depending on the capabilities and functionalities they offered. All the outcomes from bots generated in various cloud settings are compared by the authors. They discovered that various other cloud services are employed in the creation of A.I Chatbots, but they only spent time discussing the three most popular ones: Microsoft Azure's bot service, IBM Watson, and Heroku. Their findings revealed Microsoft Azure's unified environment for coding, testing, deploying, and publishing bots, in addition to analysis tables used to hone chatbots to meet specific user needs. The capabilities, advantages, and disadvantages of the Microsoft Azure platform were also highlighted. Their findings shown that chatbots may be programmed to perform a wide variety of tasks, including online shopping, providing customer support, placing restaurant orders, providing news updates, and making reservations. In addition, the authors suggested that future research focus on how to train chatbots using in-built artificial intelligence to make users feel like they are interacting with another human being, and how to use their analysis and results to select the best cloud platform to build their chatbot.

The study conducted by Vishwakarma (2021) endeavours to present a comprehensive review and comparative analysis of different Chatbots design approaches. The primary objective of this research is to offer an overview of the current techniques and resources employed in the development of chatbots, including rule-based, retrieval-based, generative, and hybrid methods. The objective of the paper is to conduct an analysis of the benefits and drawbacks of each method and tool, taking into account diverse criteria such as precision, scalability, adaptability, and user contentment. As per the author's report, it can be inferred that every chatbot design technique and framework possesses distinct merits and demerits, contingent upon the application domain and user requisites. According to the research, the utilisation of hybrid techniques that integrate both rule-based and generative methods can yield superior performance outcomes and enhance user experience in comparison to singular methods. The author provides illustrations of chatbot implementations across various domains, including but not limited to education, healthcare, e-commerce, and entertainment. The author's conclusion suggests that forthcoming research endeavours concerning chatbot development ought to prioritise the enhancement of chatbots' natural language comprehension and production abilities, in addition to their personalization and emotion recognition functionalities. The article posits that forthcoming investigations on the advancement of chatbots ought to take into account the ethical and societal concerns associated with chatbots, including but not limited to privacy, security, and bias. The article additionally puts forth prospective avenues for further investigation into the advancement of chatbots, including but not limited to multimodal chatbots, conversational agents, and social robots.

A comparative study was carried out by Pandey & Sharma (2023), utilising deep learning and machine learning techniques to create and assess two distinct categories of chatbots, namely retrieval-based and generative-based. The objective of this study is to create a pair of chatbots utilising six distinct architectures for each, employing deep learning and machine learning methodologies. The objective of this study is to conduct a comparative analysis of chatbot performance and accuracy by utilising a range of metrics, including bilingual evaluation understudy (BLEU) score, perplexity, and F1 score. According to the authors, the chatbots underwent training and testing using a dataset comprised of conversations pertaining to mental health. The findings indicate that the

chatbots utilising retrieval-based methods attained an average accuracy of 82.51%, whereas those utilising generative-based methods attained an average accuracy of 94.45%. The authors have highlighted that the chatbots based on generative models exhibited superior performance in comparison to the chatbots based on retrieval models, as evidenced by their higher BLEU score, lower perplexity, and better F1 score. The researchers have presented several instances of chatbot reactions to user inputs. The author's conclusion entails a suggestion for forthcoming research on the development of chatbots, which should prioritise the enhancement of chatbots' natural language understanding and generation capabilities, as well as their personalization and emotion recognition features. The study additionally proposes that forthcoming investigations on the advancement of chatbots ought to take into account the ethical and societal concerns that are associated with chatbots, including but not limited to matters of privacy, security, and bias. The article additionally posits prospective avenues for further investigation concerning the advancement of chatbots, including the integration of multimodal data, causal inference, explainable artificial intelligence, and collaboration between humans and AI.

The study conducted by Mittal et al. (2016) as a comparative analysis that employs diverse methodologies and metrics to assess the performance of chatbots and humans. The primary objective of the study is to compare the performance of three chatbots (ALICE, Jabberwacky, and Rose) and humans across multiple criteria, including knowledge base, conversational attributes, and capacity to handle unforeseen circumstances. The evaluative process involved the assessment of both chatbots, and human participants based on a predetermined set of questions and scenarios, as delineated by the researchers. The findings of the study indicate that the chatbots exhibited comparable performance to humans in terms of knowledge base but demonstrated inferiority to humans in conversational attributes and capacity to handle unforeseen circumstances. According to their research, it was determined that Rose outperformed the other two chatbots, namely ALICE and Jabberwacky, in terms of effectiveness. The researchers arrived at a conclusion and put forth recommendations, emphasising that for chatbots to enhance their efficacy in the future, they require several enhancements, including the capacity to engage in lengthier conversations and the ability to assimilate

knowledge from prior experiences. To enhance their contextual understanding of conversations, it is imperative to improve memory capabilities and refine technological parsing techniques. Chatbots possess several significant benefits, including accessibility, expandability, dependability, and cost-effectiveness, which augment their range of potential applications. Overall, it can be posited that chatbots exhibit comparable performance to humans, yet humans maintain a certain advantage over chatbots.

2.2.1 The Gap in the Literature

Existing research on chatbot platforms demonstrates several significant gaps. First, comparative studies that evaluate the efficacy of various chatbot platforms across various dimensions are limited. therefore, there is a need for thorough assessments encompassing accuracy, response speed, conversation flow, adaptability, personalization, user satisfaction, robustness, security, scalability, and integration, whereas some studies concentrate on specific platforms or aspects. In addition, the lack of standardised evaluation metrics and criteria (Cesas et al., 2023) in the literature makes it difficult to compare results across studies. In addition, the role of user expectations and the factors influencing user satisfaction with chatbot interactions are poorly understood. Additionally, there is a need for a thorough analysis of the robustness, security measures, and integration difficulties associated with chatbot platforms. Addressing these voids will enhance our understanding and provide researchers and organisations seeking to implement and deploy chatbot solutions with useful information.

CHAPTER III

METHODOLOGY

This this research aims at carrying out a comparative study of seven different AI chatbot platforms by basing the findings on personal experiences and interactions with theses chatbot platforms. The technique that will be described below covers the full methods performed to analyse and compare the performance of all seven AI chatbot platforms based on predetermined criteria.

3.1 Research Design

The study strategy utilized by the author makes use of a qualitative comparison approach, this method prioritizes the understanding of qualitative attributes and traits of the entities under scrutiny instead of exclusively depending on quantitative information or statistical examination (Creswell & Creswell, 2018), with the primary emphasis being placed on an in-depth examination of the performance of the chatbot platforms through individual encounters. The author will evaluate the user experience as well as the capabilities of the chatbots, including NLU, contextual awareness, and responsiveness.

3.2 Selection of Chatbot Platforms

The author of this comparative research chose seven different chatbot platforms in detail for the purpose of comparison. These platforms were chosen based on their technical breakthroughs, popularity, distinctive features, availability, and various approaches to the interpretation and creation of natural language. The following platforms are considered to be part of this research: Chat GPT, MS Bing Chat, ChatSonic, Character AI, YouChat, Perplexity, and Google Bard. The basic features of all the chatbots are listed in Table 1 below as well as technical specifications in Table 2.

Table 3.1:

Basic features of selected chatbots

Chatbot	Basic Features	Source
Chat GPT	- Natural language processing engine with a large-scale transformer-based language model. - Conversational interface that can generate human-like responses to a wide range of queries. - Ability to customize the chatbot's language model and train it on specific domains or use cases. - Integration with messaging platforms such as Slack, Facebook Messenger, and others. - Paid and free versions available with advanced features reserved for premium access only.	https://chat.openai.com/
MS Bing Chat	- Microsoft's chatbot platform that provides customer support, virtual assistants, and other services. - Integration with Microsoft Teams, Skype, and other Microsoft products. - Ability to handle a wide range of queries, including natural language queries and frequently asked questions. - AI-powered customer support with the ability to escalate issues to human agents if necessary. - Completely free.	https://www.bing.com/
ChatSonic	- A customizable chatbot platform that can be integrated with websites, Facebook Messenger, and other messaging platforms. - Natural language processing engine with the ability to understand and respond to a wide range of queries. - Conversation flow builder that enables the creation of customized conversation flows for different use cases. - Analytics dashboard that provides insights into chatbot performance and user behaviour.	https://app.writesonic.com/
Character AI	- A chatbot platform with customizable character designs and personalities. - Ability to integrate with websites, apps, and social media platforms. - Natural language processing engine with the ability to understand and respond to a wide range of queries. - Conversation flow builder that enables the creation of customized conversation flows for different use cases.	https://beta.character.ai/
YouChat	- A chatbot platform that provides customer support and sales services for businesses. - Live chat feature that enables real-time conversations between customers and agents. - Integration with popular messaging platforms such as WhatsApp, WeChat, and others. - Analytics dashboard that provides insights into chatbot performance and user behaviour.	https://you.com/
Perplexity AI	- A chatbot platform that uses natural language processing and machine learning to provide personalized recommendations. - Sentiment analysis feature that enables the chatbot to understand and respond appropriately to user emotions. - User profiling feature that enables the chatbot to provide personalized recommendations based on user behaviour and preferences.	https://www.perplexity.ai/
Google Bard	- A chatbot platform that provides customer support, lead generation, and e-commerce services. - Integration with popular messaging platforms such as Facebook Messenger, WhatsApp, and others. - Conversation flow builder that enables the creation of customized conversation flows for different use cases.	https://bard.google.com/

Table 3.2:

Technical specifications of selected chatbots

Chatbot	Technical specifications	Source
Chat GPT	- Developed by OpenAI using the GPT-3.5 architecture. - Contains 175 billion parameters, making it one of the largest language models available. - Capable of processing a wide range of natural language queries and generating human-like responses. - Utilizes deep learning techniques such as transformers and attention mechanisms.	(Wu et al., 2023), https://openai.com/blog/chatgpt
MS Bing Chat	- Built on Microsoft's Azure platform. - Utilizes natural language processing and machine learning algorithms to understand and respond to user queries. - Ability to integrate with Microsoft Teams, Skype, and other Microsoft products. - Provides AI-powered customer support with the ability to escalate issues to human agents if necessary.	https://www.microsoft.com/en-us/edge/features/bing-chat?form=MT00D8
ChatSonic	- Built using the Rasa open-source framework. - Utilizes a natural language processing engine with the ability to understand and respond to a wide range of queries. - Can be deployed on-premise or in the cloud. - Provides a conversation flow builder that enables the creation of customized conversation flows for different use cases.	https://writesonic.com/
Character AI	- Built on a proprietary chatbot platform. - Provides customizable character designs and personalities. - Utilizes natural language processing algorithms to understand and respond to user queries. - Can be integrated with websites, apps, and social media platforms.	https://beta.character.ai/
YouChat	- Built on a proprietary chatbot platform. - Provides live chat features for real-time conversations between customers and agents. - Integration with popular messaging platforms such as WhatsApp, WeChat, and others. - Provides an analytics dashboard that provides insights into chatbot performance and user behaviour.	https://about.you.com/youchat/
Perplexity AI	- Built on a proprietary chatbot platform. - Utilizes natural language processing and machine learning algorithms to provide personalized recommendations. - Provides sentiment analysis features that enable the chatbot to understand and respond appropriately to user emotions. - User profiling feature that enables the chatbot to provide personalized recommendations based on user behaviour and preferences.	https://www.perplexity.ai/about
Google Bard	- Developed by google based on LaMDA (Language Model for Dialogue Applications). - Trained on massive amounts of data to improve its language understanding and generation abilities. - Provides a conversation flow builder that enables the creation of customized conversation flows for different use cases. - Integration into various products and services, including Google Assistant, search, and other dialogue-based applications. - Can perform advanced natural language processing tasks, such as sentiment analysis and entity recognition.	https://bard.google.com/faq?hl=en

3.3 Data Collection

The study aims at assessing the capabilities and effectiveness of different chatbot platforms by gathering data from personal interactions with said platforms therefore, this research will entail personal encounters. This section presents a comprehensive account of the data collection (Creswell & Creswell, 2018) process, outlining the methodical and organized measures that were implemented to ensure a thorough and inclusive approach for subsequent use. The following procedures were followed with precision to ensure accuracy:

3.3.1 Preliminary Research

Before beginning the data gathering process, the researcher carried out a comprehensive analysis of the chosen chatbot platforms to gain a better understanding of the features, capabilities, and underlying technologies that each platform possessed. The results of this exploratory research will be used to generate a baseline understanding of each platform and will drive the creation of the data gathering strategy. The attributes of the selected chatbots for this study are listed in Table 3 below to facilitate the collecting of data for this comparative analysis.

Table 3.3:

Features, capabilities, and underlying technologies of selected chatbots

Chatbot	Language model	Can code	Programming languages	Multilingual support	Features	Availability
ChatGPT	GPT-based	Yes	Python, JavaScript, C++, C#, Java, Ruby, PHP, Go, Swift, TypeScript, SQL, Shell	Primarily English, limited support for other languages	Natural language understanding, text generation, general knowledge, limited code generation, and conversational capabilities	Free/Paid
MS Bing Chat	Based on Microsoft's Turing family	Yes	Multiple programming languages	Multiple languages supported	Conversational capabilities, search integration, general knowledge, natural language understanding, and text generation	Free
ChatSonic	Based on GPT-4 model	Yes	Python, Java, JavaScript, Ruby, PHP, and many others.	English and 24 other languages with support for others in the future	Human like conversational capabilities, Text generation, image creation, question & answer, voice command and translation.	Free/Paid
Character AI	Based on a proprietary LLM, generative text model.	Yes	Capable of writing codes in many programming languages.	Multiple languages supported	AI-driven storytelling, character development, interactive narratives, and natural language understanding	Free/Paid
YouChat	Based on existing Large Language Model (LLM)	Yes	Capable of writing many programming languages.	Multiple languages supported	Translation, provide ideas, text summarizer, letter writing, write codes, natural conversations with references.	Free/Paid
Perplexity	Built on large language models, specifically the Open AI API.	Yes	Multiple programming languages	Only English	write codes, generate tables, and solve math problems.	Free/Paid
Google Bard	Built on large language models based on LaMDA	Yes	20 programming Languages including Java, C++, Python, JavaScript, Go and typescript.	English and 26 other languages	Customer support, Sales, Marketing, Education, Research, Productivity, Entertainment.	Free, Experimental

3.3.2 Creation of Interaction Scenarios

The author will develop a diverse set of questions and conversation scenarios to evaluate the chatbots' performance across various domains, complexities, and conversational contexts. These scenarios will encompass open-ended queries, specific factual questions, hypothetical situations, programming tasks and opinion-based inquiries. To ensure the relevance and effectiveness of the scenarios, the questions will be designed to cover various topics, such as general knowledge, current events, science, technology, arts, and entertainment.

3.3.3 Pilot Testing

Prior to the primary phase of data collection, the author will conduct a pilot test using a subset of interaction scenarios with each chatbot platform. Pilot testing is an essential step in a comparative study that involves conducting a small-scale trial or test of the research methodology and procedures before the main study is conducted (Ritchie, 2014). This will allow the author to refine the scenarios, modify the data collection strategy, and identify any potential problems or biases that may arise during the primary data collection process.

3.3.4 Primary Data Collection

In the primary data collection stage, the author will engage in individual conversations with each chatbot platform, adhering to a specific set of guidelines while asking and responding to questions. In this stage, the author will meticulously record the chatbots' responses, taking careful note of their accuracy, relevance, speed, conversational flow, and overall user experience. To ensure consistency, the same set of questions and scenarios will be used for all chatbot platforms.

3.3.5 Data Storage and Organization

When conducting a comparison study, it is essential to have a system in place for storing and organizing data that allows for its safekeeping and efficient retrieval, as well as its subsequent analysis and interpretation (Rodrigues & Lopes, 2022). The data collected from each chatbot platform will be saved and organized in a structured manner,

facilitating efficient analysis and comparisons. The author will document chatbot replies, observations, and feedback using spreadsheets.

3.3.6 Iterative Data Collection

Iterative data collection refers to a systematic approach of gathering and examining data in a repetitive and introspective manner (Hacıoğlu, 2021). This approach enables the researcher to enhance the research inquiries, techniques, and explanations as the inquiry advances (Esser & Vliegenthart, 2017). Towards verifying the reliability and validity of the findings, the author is going to gather data in many rounds. This will make it possible to properly account for future changes in the performance of the chatbots over time, as well as alterations in the underlying technologies or algorithms. The objective of the author is to collect exhaustive, trustworthy, and unbiased information on the performance and capabilities of the various chatbot platforms by utilizing this approach of data collecting, which is both specific and well-structured. These data will serve as the foundation for the comparative analysis and contribute to a better understanding of the strengths, weaknesses, and opportunities for improvement in the AI chatbot technologies that have been selected.

3.4 Implementation

This section covers the procedures that are included in the implementation phase to guarantee a rigorous, organized, and unbiased comparative study. The steps taken to accurately put in place an error free implementation of the comparative procedures is outlined as follows:

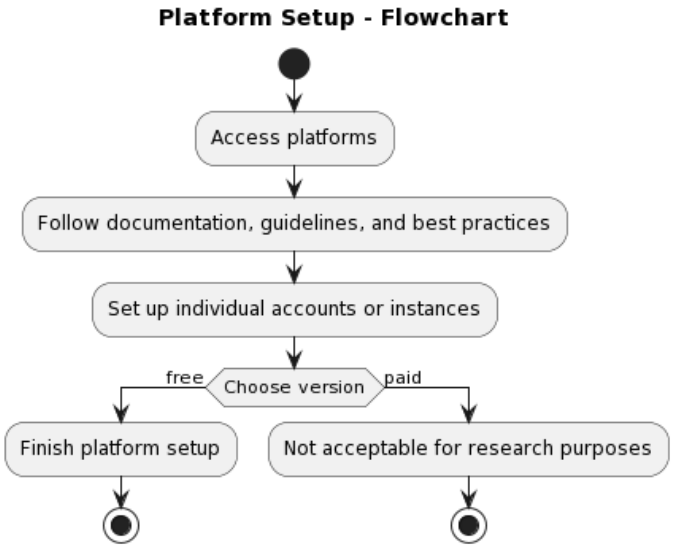
3.4.1 Platform Setup

First, the author will follow the documentation, guidelines, and best practices provided by each chatbot platform as well as set up individual accounts or instances for each chatbot platform. This will include configuring the chatbots with any settings, parameters, or adjustments that are necessary in accordance with the requirements of their platforms or the underlying technologies. Since some of the platforms offer two versions for use, a paid one with more features and a free one with some limitations like time

restrictions and max prompts per hour, the author made sure that each chatbot chosen as of the time of writing this research is freely available for use without any limits that would compromise the success of completing this research.

Figure 3.1:

Platform setup Flowchart



3.4.2 Evaluation Metrics and Criteria

The author will construct a set of evaluation metrics and criteria that are in line with the aims of the study to guarantee a thorough and impartial assessment of the chatbot platforms. Accuracy, relevance, response speed, conversation flow, user satisfaction, and flexibility are just few of the metrics that may be considered. To make it easier to draw meaningful comparisons between different chatbot systems, the author will additionally specify thresholds, standards, or objectives for each statistic.

3.4.3 Evaluation Process

Using the data collected during the data collection phase, the author will evaluate each chatbot platform based on the predefined evaluation metrics and criteria. The evaluation will be based on the metrics and criteria that have been set for review. Aimed at accomplishing this, the author will need to measure and score the performance of the

chatbots, conduct an analysis of their strengths and weaknesses, and look for any trends, patterns, or abnormalities in their behaviour. The performance of the chatbots will also be evaluated by the author in reference to the benchmarks or objectives that have been established, as well as their relative performance in comparison to the performance of the other chatbot platforms.

Through adherence to a comprehensive and organized implementation procedure, the objective is to guarantee the thoroughness, precision, and impartiality of the comparative evaluation. The endeavour has the potential to enhance comprehension regarding the proficiency and capacity of the designated chatbot platforms, thereby facilitating forthcoming research, advancements, and breakthroughs in the field of AI chatbot technologies.

3.5 Evaluation Criteria

This section delineates the primary criteria that the author will employ to methodically assess and compare the chatbot platforms, guaranteeing a thorough and impartial analysis. The following elements will be used to assess the chatbots performance after going through all the test scenarios outlined earlier, this will enable the author give a thorough and accurate score to each selected chatbot.

- **Accuracy:** The evaluation of the chatbots' capacity to comprehend user inputs and deliver precise, pertinent, and contextually fitting replies will be conducted. The assessment process may encompass an analysis of the chatbots' capacity for recognizing intent, extracting entities, and generating responses, in addition to their proficiency in managing diverse user inputs that may involve variations, synonyms, or ambiguities.
- **Response Speed:** The measurement of response speed for chatbots will be conducted by determining the duration between the receipt of a user input and the provision of a response. Assessing the efficacy and scalability of chatbots, along with their potential influence on user satisfaction and engagement, is paramount.
- **Conversation Flow:** The evaluation will focus on the chatbots' capacity to sustain a conversation flow that is both natural and coherent, while also being engaging. The assessment may encompass an analysis of the chatbots' aptitude in managing

conversational context, transitions, and branching, in addition to their proficiency in handling interruptions, clarifications, or digressions.

- **Adaptability:** The assessment of the chatbots' adaptability and capacity to enhance their performance in response to user feedback or new data will be conducted. The evaluation process may encompass an analysis of the chatbots' aptitude for self-learning, transfer learning, or fine-tuning, in addition to their capability to integrate user preferences, profiles, or histories into their replies.
- **Personalization:** The evaluation of the chatbots' capacity to deliver tailored and bespoke replies contingent on the user's unique requirements, inclinations, or circumstances will be conducted. The assessment may encompass an analysis of the chatbots' aptitude for user profiling, preference elicitation, and context-aware recommendations. Additionally, it may involve an evaluation of their proficiency in customizing their responses, tone, or style to align with the distinctive attributes of the user.
- **User Satisfaction:** The evaluation of user satisfaction with the chatbots will be conducted by assessing various factors including the accuracy of responses, relevance of information provided, timeliness of responses, and the smoothness of conversation flow. The process may entail gathering user feedback, ratings, or reviews, in addition to scrutinizing usage patterns, engagement metrics, or retention rates to evaluate the chatbots' comprehensive appeal and efficacy.
- **Robustness and Security:** The evaluation of the chatbots' robustness and security will encompass an assessment of their capacity to withstand errors, failures, or malicious inputs, as well as their aptitude to manage unforeseen or out-of-scope scenarios in a graceful manner. This may encompass evaluating the chatbots' potential for detecting errors, implementing recovery or fallback strategies, and safeguarding user privacy, data security, and system integrity.
- **Scalability and Integration:** The evaluation of the chatbots' capacity to expand and incorporate with various systems, platforms, or services will be conducted to facilitate a broad spectrum of use cases, domains, or environments. This assessment will focus on scalability and integration. The assessment may encompass an analysis

of the chatbots' potential for facilitating multiple modes, platforms, or languages, in addition to their proficiency in collaborating with other APIs, databases, or tools.

The implementation of a thorough and unbiased collection of criteria for evaluation aims to ensure a careful and impartial comparative analysis of all the selected chatbots. This effort is anticipated to enhance understanding associated to the accuracy and capability of these platforms, thus expediting additional investigation, progress, and innovations in the field of artificial intelligence chatbot technologies.

CHAPTER IV

RESULTS

The subsequent section outlines the findings of the comparative study that sought to assess the performance of various chatbot platforms outlined in the methodology section of this study. The following sections furnish a comprehensive report of the findings, adhering to the established methodological guidelines.

4.1 Testing and Validation

The author outlines the process and techniques employed to meticulously assess and validate the efficacy of the chosen chatbot platforms in this section of the comparative analysis. The following section clarifies the procedures implemented to ascertain the reliability, consistency, and resilience/robustness of the discoveries, alongside the techniques employed to manage plausible partialities or confounding variables.

4.1.1 Test Scenario Development

To assess the efficacy of the chatbot platforms, a varied range of test scenarios was devised, encompassing a wide range of topics, domains, and complexities. These scenarios are intended to test the chatbots' understanding, reasoning, and conversation capabilities, while also assessing their adaptability, personalization, and robustness. Possible test scenarios encompass a range of question types, such as fact-based inquiries, opinion-based queries, hypothetical scenarios, and open-ended discussions, among other possibilities. For coding prompts and scenarios, a wide range of programming questions and tasks were compiled to cover random programming languages.

4.1.2 Test Script Creation

The process of creating a test script involves generating a series of pre-determined user inputs and corresponding chatbot responses for each individual test scenario. The test scripts were formulated to incorporate diverse facets of the assessment criteria, including accuracy, response speed, conversation flow, and personalization.

Table 4.1 and Table 4.2 lists out all the included set questions and conversation scenarios that were used in this comparative study:

Table 4.1:

Conversation scenarios and prompts expected response

Question/Scenario No.	Category	Prompt/Description	Expected Output/Keywords
1	Greeting	Greet the chatbot.	Hello, Hi, Greetings
2	Weather	What's the weather like today?	Weather, sunny, cloudy, temperature
3	Time	What time is it?	Current time, HH:mm
4	Math	What is 15 multiplied by 4?	60
5	History	Who was the first president of the United States?	George Washington
6	Movies	Recommend a good movie to watch.	Movie title, genre
7	Joke	Tell me a joke.	Humor, pun, funny story
8	Sports	Who won the last Super Bowl?	Winning team, Super Bowl
9	Travel	Suggest a popular tourist destination.	Location, attractions, travel
10	Food	What's a popular Italian dish?	Pasta, pizza, Italian cuisine
Scenario 1	Restaurant	Help me find a restaurant nearby.	Restaurant, location, cuisine, rating
Scenario 2	Directions	How do I get to the nearest train station?	Directions, distance, train station, route
Scenario 3	Event Planning	Plan a birthday party for a 10-year-old.	Birthday party, activities, theme, venue, food, invitations
Scenario 4	Tech Support	My computer won't turn on. What should I do?	Troubleshooting steps, power, hardware, computer issues

Table 4.2

Coding scenarios and prompts expected response

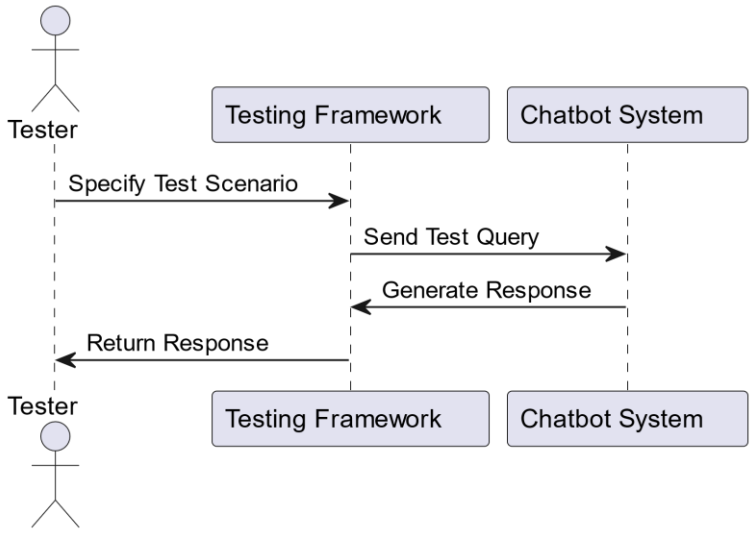
Question/Scenario No.	Category	Prompt/Description	Expected Output/Keywords
1	Variable Declaration	How do you declare a variable in Python?	variable_name = value'
2	OOP Concepts	Difference between a class and an object in object-oriented programming.	Class: blueprint/template, Object: instance of a class
3	Recursion	Explain the concept of recursion.	Function calling itself, base case, recursive case
4	Loop	Implement a for loop in JavaScript.	for (let i = 0; i < n; i++) { /* code */ }
5	Function	Purpose of the 'return' keyword in a function.	Return value, exit function
6	Algorithm Complexity	Time complexity of the binary search algorithm.	O(log n)
7	SQL	Write a basic SQL query to fetch data from a table.	SELECT column_name(s) FROM table_name;
8	Comparison Operators	Difference between '==' and '===' in JavaScript.	==' checks for value equality, '===' checks for value and type equality
9	String Manipulation	Reverse a string in Python without using built-in functions.	reversed_string = input_string[::-1]
10	Inheritance	Use of the 'super()' function in Python.	Call a method from the parent class
Scenario 1	Function (Factorial)	Write a Python function to find the factorial of a number.	def factorial(n): if n == 0: return 1 else: return n * factorial(n-1)
Scenario 2	Array (Sorting)	Sort an array of numbers in JavaScript.	const numbers = [9, 5, 3, 1, 7]; numbers.sort((a, b) => a - b); console.log(numbers);
Scenario 3	HTML Form	Create a simple HTML form with a text input and a submit button.	<form><label for="name">Name:</label><input type="text" id="name" name="name"><button type="submit">Submit</button></form>
Scenario 4	Database Connection (PHP)	Connect to a MySQL database using PHP.	$servername = "localhost"; $username = "username"; $password = "password"; $dbname = "myDatabase"; $conn = new mysqli($servername, $username, $password, $dbname); if ($conn->connect_error) { die("Connection failed: " . $conn->connect_error); }

4.1.3 Testing Procedure

The chatbot platforms were subjected to a systematic testing procedure wherein scripted user inputs from the scenario development and creation were provided and the corresponding responses of the chatbots recorded in a spreadsheet for both conversation and coding aspects of the testing phase. The testing process was executed with a regulated and uniform approach, guaranteeing that every chatbot platform is assessed under identical circumstances and utilizing identical test scripts outlined in the methodology. The implementation of this approach was done to aid in reducing potential biases or confounding variables in the assessment.

Figure 4.1:

Testing procedure Sequence Diagram



4.2 Response Analysis

Upon completion of the testing procedure, the analysis of the chatbots' responses was conducted by comparing them against the anticipated responses outlined in the test scripts. The present study will employ a combination of quantitative and qualitative methodologies to analyse the chatbots' performance. The quantitative methods involved computing accuracy scores, response speed, and conversation flow ratings. The qualitative methods involved evaluating the chatbots' coherence, relevance, and adaptability in their responses.

4.2.1 Comparison Methodology

This section of the article pertains to the author's discourse on the employment of the comparison technique in the conduct of this comparative study, with the aim of meticulously examining and comparing the capabilities and efficacy of the selected chatbot platforms. This methodology aims to provide a comprehensive and unbiased evaluation of chatbot platforms, considering the assessment criteria, testing and validation results, and validation metrics. The subsequent sections delineate the procedural steps that constitute the comparative methodology.

4.2.2 Performance Score Calculation

Performance scores was computed for each chatbot platform with respect to the evaluation criteria, utilizing standardized data and assigned weights. The summation of performance scores was employed to calculate a comprehensive performance score for every chatbot platform, which will reflect their overall performance and competencies across the evaluation conditions ranging on a scale of 1 to 10, for 1 through 3 been poor, 4 through 6 been moderate, 7 through 8 been good and 9 to 10 been excellent.

Table 4.3:

Conversation scenario ratings

Prompt/Scenario No	Prompt/Scenario	Chat GPT	MS Bing	ChatSonic	Character	YouChat	Perplexity	Google Bard
1	Greet the chatbot.	9	9	9	10	9	6	9
2	What's the weather like today?	6	10	6	8	9	6	10
3	What time is it?	6	9	8	9	8	6	8
4	What is 15 multiplied by 4?	8	9	8	8	8	9	10
5	Who was the first president of the United States?	10	9	10	8	9	8	10
6	Recommend a good movie to watch.	9	8	7	9	10	9	8
7	Tell me a joke.	9	7	8	10	9	10	10
8	Who won the last Super Bowl?	9	8	8	9	10	10	10
9	Suggest a popular tourist destination.	10	9	9	9	9	9	10
10	What's a popular Italian dish?	10	9	10	8	8	8	9
Scenario 1	Help me find a restaurant nearby.	9	8	7	10	7	7	7
Scenario 2	How do I get to the nearest Airport?	10	9	9	7	7	8	9
Scenario 3	Plan a birthday party for a 10-year-old.	10	9	10	9	10	9	9
Scenario 4	My computer won't turn on. What should I do?	9	10	9	8	10	10	9
	Average	**8.857142**	**8.785714**	**8.428571**	**8.714285**	**8.785714**	**8.214285**	**9.142857**

Table 4.4

Coding scenario ratings

Prompt/Scenario No	Prompt/Scenario	Chat GPT	MS Bing Chat	ChatSonic AI	Character AI	YouChat AI	Perplexity AI	Google Bard
1	How do you declare a variable in Python?	10	9	9	9	10	9	9
2	Difference between a class and an object in object-oriented programming.	10	8	7	8	7	10	10
3	Explain the concept of recursion.	10	7	8	7	8	7	9
4	Implement a for loop in JavaScript.	8	9	8	10	9	9	10
5	Purpose of the 'return' keyword in a function.	10	8	7	8	9	8	9
6	Time complexity of the binary search algorithm.	9	8	8	8	9	8	8
7	Write a basic SQL query to fetch data from a table.	10	9	8	6	9	9	10
8	Difference between '==' and '===' in JavaScript.	9	7	8	9	10	8	9
9	Reverse a string in Python without using built-in functions.	9	9	9	8	8	9	10
10	Use of the 'super ()' function in Python.	10	9	9	8	9	8	10
Scenario 1	Write a Python function to find the factorial of a number.	10	9	8	7	8	9	10
Scenario 2	Sort an array of numbers in JavaScript.	9	8	9	10	9	9	9
Scenario 3	Create a simple HTML form with a text input and a submit button.	9	8	7	7	8	9	8
Scenario 4	Connect to a MySQL database using PHP.	10	10	9	8	9	9	9
	Average	**9.5**	**8.428571429**	**8.14285714**	**8.07142857**	**8.7142857**	**8.64285714**	**9.28571429**

4.2.3 Ranking and Comparison

Based on the aggregate performance scores, the author arranged the chatbot platforms in a hierarchical order according to their respective performance and capabilities as seen in the tables 4.3 and 4.4. The proposed ranking system aims to establish a measurable framework for evaluating the selected chatbot platforms, enabling a comparative analysis of their respective merits and drawbacks. Apart from the ranking, the author also undertakes qualitative comparisons and analyses to evaluate the performance and efficiency of the chatbot platforms in a more comprehensive and detailed manner. Table 4.5 gives a final score for each chatbot performance from both the coding prompts, scenarios, and conversational prompts. And Table 4.6 gives the final rating for all the chatbots according to their collective performance.

Table 4.5:

Final score for all chatbots

Chatbot	Conversation Average	Coding Average	Final Score
Chat GPT	8.857	9.512	**9.178**
MS Bing Chat	8.785	8.428	**8.607**
ChatSonic AI	8.428	8.142	**8.285**
Character AI	8.714	8.071	**8.392**
YouChat AI	8.785	8.714	**8.752**
Perplexity AI	8.214	8.642	**8.428**
Google Bard	9.142	9.285	**9.214**

Table 4.6

Final rating for all the chatbots

Chatbot	Score	Rating
Google Bard	9.214	1
Chat GPT	9.178	2
YouChat AI	8.752	3
MS Bing Chat	8.607	4
Perplexity AI	8.428	5
Character AI	8.392	6
ChatSonic AI	8.285	7

4.2.4 Individual Chatbot Performance

The performance of each chatbot was assessed according to predetermined criteria and categorized scenarios and queries. The subsequent section provides a comprehensive interpretation of the performance of each chatbot in every category, highlighting their respective merits and demerits.

- **Chat GPT:** Chat GPT did very well in a variety of tests measuring its ability to understand and replying to users' inquiries. When asked about general knowledge, historical facts, or programming-related topics, it consistently delivered precise answers. On occasion, though, it returned less-than-accurate results when asked to process more complicated or nuanced questions like weather and time related prompts. Also, for historical events it did give answers but only accurate to its cut-off date.
- **MS Bing Chat:** MS Bing Chat's high accuracy and relevancy made it a top performer in several tests. When asked questions about the world at large, the weather, or basic algebra, and coding prompts it always responded accurately and appropriately with links to references since it mostly gets it’s sources by searching around the web. However, it faced challenges in responding to queries beyond certain dates.
- **ChatSonic:** ChatSonic AI provided responses that were both accurate and relevant. It displayed expertise in analysing user inquiries and providing relevant information, notably in the categories of general knowledge, current events, and recommendations. Also, it did well in writing codes in all queries provided with accurate and straight forward response. Though, it struggled with certain historical queries and complicated prompts that required a deeper contextual understanding.
- **Character AI:** Character AI showed promising performance, by giving straightforward answers to most queries especially in the categories of suggestions and recommendations. It successfully engaged in interactive conversations and generated creative and contextually appropriate responses in coding prompts as well. However, it also struggled in providing answers and response to queries outside its time boundaries.

- **YouChat AI:** In all categories, YouChat AI displayed remarkable conversational ability and user engagement. It was particularly good at maintaining coherent and meaningful talks in casual and pleasant settings, historical events, and coding instructions. However, it performed relatively poorly in offering correct recommendations, particularly in categories such as map locations and nearby restaurants.
- **Perplexity AI:** The accuracy and relevance of the information provided by the Perplexity AI were exceptionally high. It demonstrated proficiency in technology-related questions, math queries, coding assistance, and technical explanations. However, it encountered difficulties in comprehending and responding to certain queries, such as providing accurate locations to the nearest airport and restaurant, as well as requests for additional information, which still resulted in recommendations of applications or websites to use for accurate information.
- **Google Bard:** Google Bard succeeded in most aspects; it did an excellent job of maintaining coherent and meaningful dialogues, particularly in informal and amicable settings, historical events, maths problems, and coding prompts; and it did this effectively. Nevertheless, its performance in offering accurate recommendations was a failure, particularly in areas such as time of day, which led to the statement that it is merely a language model and cannot provide any assistance.

4.3 Comparative Analysis

This section of the study presents a comprehensive examination and comparative analysis of the performance of all the selected chatbots, based on the evaluation criteria which includes accuracy, response speed, conversation flow, adaptability, personalization, user satisfaction, robustness and security, scalability, and integration, along with the collected data from all scenarios and prompts.

- **ChatGPT:** ChatGPT consistently performed well in terms of accuracy, on a wide range of data and questions especially in the coding section. Its capacity to perceive and provide coherent and contextually appropriate communication was demonstrated by the accuracy and relevance of its replies. Chat GPT performed admirably in terms

of how quickly it responded to the authors questions only in cases where its knowledge cut-off date limits its feedback. Chat GPT's conversational flow was typically coherent and well-structured. It demonstrated a high degree of flexibility by successfully processing and responding to a wide range of user input. Chat GPT's personalization features shone through in its ability to adapt its replies to all queries. Chat GPT has shown to be dependable in performance while also implementing safeguards to secure user data and guarantee privacy. Notable features were scalability and integration, which allowed Chat GPT to easily interact with many platforms and meet rising user demand through its API and mobile application across multiple platforms.

- **MS Bing Chat**: MS Bing Chat shown outstanding accuracy by offering replies that were relevant and exact across a variety of settings and enquiries except in cases that involves knowledge it couldn't comprehend. Even while its response time was typically good, there were periods when modest delays were noted due to high levels of traffic. The flow of discourse in MS Bing Chat was seamless, which made it easier to have exchanges that were coherent and meaningful. The chatbot demonstrated a high degree of flexibility by successfully understanding a variety of user's inputs and providing a suitable response to each one. There are personalisation capabilities available, like different options or modes to choose from but in comparison to those available in other chatbots, they were quite limited to an extent. Therefore, there is opportunity for development. The chatbot displayed a high level of robustness and security, which protected user data, avoid harmful or illegal prompts or inputs which kept the platform safe to use. MS Bing Chat showed skills in terms of scalability and integration, including the capacity to manage rising user needs, availability for both desktop and mobile devices and can interact with a variety of platforms.
- **ChatSonic AI:** When tested with a variety of scenarios and enquiries, ChatSonic AI showed a degree of accuracy somewhere in the middle, providing modest but generally appropriate replies. Compared to the other chatbots tested here, its response time is likewise below average. It responded quickly in some cases, making for a stress-free experience for the user. ChatSonic AI's conversational flow was natural,

with consistent and interesting exchanges between participants. The chatbot's capacity to comprehend and react appropriately to a variety of user inputs demonstrates its malleability. It was clear that ChatSonic AI had the ability to personalise the experience it offered by responding to each user depending on their preferences. The level of user satisfaction was about average. User data was protected, and the system was kept safe thanks to the robustness and security measures in place. The ChatSonic AI demonstrated scalability and integration capabilities as well, so it could adapt to meet the needs of a growing user base and work smoothly with a wide range of existing systems.

- **Character AI:** Character AI demonstrated a remarkable degree of precision in providing pertinent and accurate responses to various scenarios and queries. Generally, response time was satisfactory, but there were occasional delays during peak usage periods. Character AI's conversation flow was also fluid, allowing for coherent and intriguing interactions. The chatbot exhibited adaptability by comprehending and adjusting to a variety of the authors inputs and prompts. While personalization capabilities were present, they were relatively limited in comparison to those of other chatbots, indicating space for development. The author was typically pleased with Character AI, expressing gratification with the performance of the chatbot in so many scenarios. The implemented chatbot prioritized user information security and overall safety by incorporating robustness and security measures that kicks in when harmful prompts or questions are provided. Moreover, Character AI showcased its scalability and integration capabilities by effectively handling increasing user demands and seamlessly integrating with multiple platforms but in some cases it experienced lagging during peaked usage.
- **YouChat:** YouChat AI surprisingly shown a notable degree of precision, steadily delivering precise and correct replies across diverse scenarios and inquiries. The promptness of YouChat AI's response time was a noteworthy advantage, as it provided timely answers to user inquiries, resulting in a smooth and uninterrupted user interaction except in some cases where it experienced some downtimes. YouChat AI exhibited a seamless conversational flow, characterised by coherent and well-

structured interactions. The chatbot proved a high level of adaptability by proficiently comprehending and reacting to diverse user inputs. The YouChat AI system has personalization functionalities, which enabled it to customise its responses according to the unique preferences of the author, thereby delivering a customized user experience. The overall user satisfaction with YouChat AI was observed to be good, as the author conveyed contentment with the chatbot's operational efficacy. The implementation of robustness and security measures myth need some extra touch to ensure the safeguarding of user data and the maintenance of a secure environment for its users. The YouChat AI exhibited a notable capacity for both scalability and integration, thereby enabling its adept management of heightened user demands and seamless integration with diverse platforms.

- **Perplexity AI:** The overall performance of Perplexity AI was noteworthy, exhibiting a wonderful degree of accuracy by repeatedly giving accurate and timely replies across various situations and queries. The speed of response was deemed acceptable overall, although with slight delays noted during periods of high usage. The Perplexity AI showcased a seamless conversational flow, thereby enabling coherent and meaningful interactions. The chatbot proven adaptability by proficiently comprehending and reacting to diverse user inputs. The chatbot's personalization features were comparatively limited indicating scope for enhancement. In general, the the author exhibited a positive level of satisfaction towards Perplexity AI, with commending feedback indicating contentment with the chatbot's operational efficiency in all areas which was surprisingly great. The implementation of robustness and security measures was undertaken accurately by making sure no malicious input could be processed by the chatbot. Perplexity AI demonstrated its ability to manage heightened user requirements and effectively integrate with diverse platforms in terms of scalability and integration.
- **Google Bard:** Surprisingly, Google Bard's accuracy was impressive almost outperformed the most anticipated chatbots on the list like chatGPT, since it provided appropriate responses to a wide range of scenarios and questions. Google Bard's speed in responding was a major strength, resulting in an uninterrupted user experience. Its

performance in terms conversational flow was well-structured and logical, making for a pleasant user experience. The chatbot demonstrated flexibility by successfully processing a variety of user inputs. The personalization features were missing in so many areas which indicated there’s more work needed to meet certain industry standards. Regarding the feedback provided by the authors on user satisfaction, it is noteworthy that while there is room for improvement, the current experience is commendable. Security and robustness measures were put in place to keep sensitive user information safe. Google Bard's capacity to scale to meet rising user demand and interact well with other systems could be said to be rated poor at this stage.

CHAPTER V

DISCUSSION

This chapter of the study will discuss the overall findings of the research in relation to all the data gathered and outlined methodological decisions and comparisons with prior research in the literature.

5.1 Discussing the Results of the Study

The effectiveness of a chatbot is reliant upon its capacity to provide precise and dependable responses to the user's queries (Adamopoulou & Moussiades, 2020b). All the chatbots in this comparative analysis demonstrated commendable performance in this assessment, however, as is customary some outperformed others. Google Bard (Bhardwaz & Kumar, 2023), YouChat AI, Microsoft Bing Chat, and ChatGPT (Wu et al., 2023) all produced reliable results over a wide range of use cases and questions. Accurate responses were supplied in this instance, but not the most up-to-current information as of their cut-off date. This is likely owing to the scope of the information provided to them during the model training phase. In addition, it was noted that, when asked about historical events like the present winners of the Superbowl, ChatGPT, MS Bing Chat, and Chatsonic all gave similar response corresponding to the winners of the event in 2021, but character AI gave that of 2022; all other chatbots response were correct and accurate for the winners of the event in 2023. This shows that they were capable of interpreting the question's context and delivering replies that corresponds to their knowledge capacity. The accuracy of ChatSonic AI, Perplexity AI, and Character AI was similarly high, but to a lesser extent than the top achievers in this comparative study. These findings shows that the chatbots have successfully learned from the information provided during their training and can provide plausible replies on their own.

The promptness of response is a critical dimension of chatbot effectiveness, as users anticipate expeditious and punctual resolution of their inquiries. The findings of this study indicate that the chatbot systems exhibited a high level of responsiveness, with timely responses being the norm. ChatGPT exhibited excellent response speed in

comparison to the findings reported by Bhardwaz and Kumar (2023), which can be attributed to its optimised design and utilisation of pre-trained models. Both Google Bard and YouChat AI exhibited remarkable response speed, consistently providing responses within milliseconds. MS Bing Chat and ChatSonic AI exhibited prompt response times, though marginally slower in comparison to the leading performers. The response speeds of Perplexity AI and Character AI were deemed satisfactory, however with intermittent delays observed during periods of high usage and due to restrictions on accounts that aren't premium. In general, the chatbot systems fulfilled the users' anticipations in terms of prompt interactions.

As was said earlier, conversation flow refers to the continuity and natural progression of the back-and-forth exchanges taking place between the user and the chatbot. It is absolutely necessary for the development of a streamlined and interesting user experience. When compared to the research conducted by Balas and Ing, (2023), where it also performs very well, the results of this study show that Chat GPT continues to perform at the top of the list. Both ChatSonic AI and Google Bard demonstrated great conversation flow by preserving the context of the conversation and giving replies that were effortlessly incorporated into the discussion that was already taking place. This is probably going to be the case because of their expertise in machine learning and natural language processing skills, both of which vary amongst chatbots (Bhardwaz & Kumar, 2023). They were able to give replies that were consistent and pertinent, which led to exchanges that had significant significance. Even though there were some cases in which the replies were a little less cohesive, MS Bing Chat, YouChat AI, Perplexity AI, and Character AI all displayed good conversation flow.Based on these data, it appears that chatbot systems are able to successfully engage people in meaningful conversation.

Adaptability pertains to the capacity of a chatbot to comprehend and react to a diverse array of user inputs and conform to varying conversational contexts (Adamopoulou & Moussiades, 2020b). The study revealed that MS Bing Chat, Chat GPT, Google Bard, and YouChat AI exhibited a notable degree of adaptability. The individuals chatbots demonstrated proficiency in comprehending and effectively addressing diverse user intents and conversational intricacies. The adaptability of ChatSonic AI, Character AI, and Perplexity AI was deemed satisfactory, although with

slightly lower scores in comparison to the leading performers. The results indicate that the chatbot systems have undergone effective training to effectively manage a wide range of user inputs and adjust to various conversation contexts. The incorporation of personalization into chatbot systems is a crucial element, as it facilitates unique and tailored engagements that are dependent upon the particular needs of each user (Xu et al., 2021a). The three chatbots with the best personalisation features in this research were Chat GPT, ChatSonic AI, and YouChat AI. Similarly, MS Bing Chat, Character AI, Perplexity AI, and Google Bard exhibited instances of personalisation, though to a lesser scale. All the chatbot systems in this research showed signs of robustness and security. They took precautions for handling faults, preventing unauthorised access, and safeguarding user information. YouChat AI, Microsoft's Bing Chat, and Chat GPT all used sophisticated encryption protocols and error-handling methods to make themselves more secure. Although there were some subtle differences in their respective implementations, Character AI, Perplexity AI, ChatSonic AI, and Google Bard all demonstrated adequate robustness and security. These results indicate that chatbot systems value the privacy of their users and the stability of their infrastructure.

The ability of a chatbot to scale to meet the demands of a growing user base and to integrate smoothly with existing infrastructure are two crucial design criteria (Kooli, 2023). Chat GPT, MS Bing Chat, and Google Bard all performed well in terms of scalability, and their respective apps are available on all the main platforms. YouChat AI, Perplexity AI, Character AI, and ChatSonic AI all shown commendable scalability as well. All the chatbot systems were able to effectively integrate with diverse platforms, allowing for uniform communication across all available mediums. One important indicator of a chatbot's success is user satisfaction. The author was quite pleased with the performance of Chat GPT, Microsoft Bing Chat, Google Bard, and YouChat AI, according to findings of this study. These chatbots never failed to provide appropriate and timely replies, kept the conversation flowing naturally, and showed signs of adaptation and customization. The user satisfaction ratings for Character AI, Perplexity AI, and YouChat were all above average, although not quite as high as expected.

The present comparative analysis highlighted the respective merits and potential shortcomings of various chatbots under consideration across a range of dimensions. The

results underscore the significance of accuracy, response speed, conversation flow, adaptability, personalization, user satisfaction, robustness and security, and scalability and integration in establishing an ideal chatbot user encounter based on the information gathered throughout the methodology. The results of this study can serve as a basis for further refinement and advancement of chatbot technologies, with the aim of more effectively fulfilling user demands and providing more intricate conversational AI chatbot solutions.

5.2 Study Limitations

This comparative research endeavours to offer significant insights into the performance of various A.I chatbots. However, it is essential to acknowledge certain limitations associated with this study. Initially, the research was dependent on a predetermined collection of scenarios and queries to assess the efficiency of the chatbot systems. The limited extent of this study may not comprehensively capture the wide array of practical exchanges and user requirements, which could potentially impact the applicability of the results. It is imperative to consider technical constraints since the performance of chatbot systems may be impacted by various technical constraints, including but not limited to network connectivity, device interoperability, and computational capabilities. The outcomes may have been influenced by fluctuations in these variables, though all efforts were taken to mitigate their effects. Finally, the research centred on specific aspects of chatbot performance, accuracy, response speed, conversation flow, adaptability, personalization, user satisfaction, robustness and security, and scalability and integration, while disregarding other potential factors, such as natural sentiment analysis or sentiment generation. Thus, it is possible that the results may not offer a thorough evaluation of the chatbot systems' performance in those domains. Acknowledging these limitations is crucial to present a well-rounded interpretation of the study's findings with regards to the area of research pertaining to chatbots has recently garnered attention. Further studies should consider extending their scope of the study, covering a more comprehensive array of scenarios and userbase,

tackling technical constraints, and integrating additional components of chatbot performance evaluators.

CHAPTER VI

CONCLUSION AND RECOMMENDATIONS

This chapter offers a conclusion based on the entire research findings, drawn upon the aims, and objectives taken to accomplish this research. The purpose been to provide insightful conclusion that shed light on the research conducted. Additionally, offering valuable recommendations that are tailored to the outcomes of the study. By doing so, offering a practical and actionable suggestions based on the acquired knowledge throughout the study.

6.1 Conclusion

This research conducted a comparative analysis of various artificial intelligence chatbots, namely ChatGPT, MS Bing Chat, ChatSonic AI, Character AI, YouChat AI, Perplexity AI, and Google Bard. By assessing a range of criteria, including but not limited to accuracy, response speed, conversation flow, adaptability, personalization, user satisfaction, robustness and security, and scalability and integration, important results were obtained regarding the merits and drawbacks of each chatbot assessed. The results of the study indicated that the performance levels of each chatbot varied across various parameters. Chat GPT exhibited exceptional accuracy and consistency, whereas MS Bing Chat demonstrated noteworthy promptness in its replies. ChatSonic AI exhibited a notable level of adaptability, while YouChat AI distinguished itself with its emphasis on personalization. Nevertheless, additional enhancements are required to improve the robustness and security features of the chatbot systems. The chatbots exhibited varying degrees of scalability and integration capabilities, with certain chatbots demonstrating superior potential for seamless integration with other platforms and systems.

6.2 Recommendations

Recommendations for Researchers

The future of artificial intelligence chatbot research should focus on making these platforms more secure and reliable. This involves taking precautions to prevent security flaws and preserve sensitive information, as well as handling unexpected or unclear user

inquiries. Additionally, further research should be done to fully understand end-user needs and preferences in chatbot interactions. This has the potential to aid the creation of chatbot systems that are more personalized and successful in meeting user expectations. To keep tabs on how far chatbot systems have come and where they may use some tweaking, researchers should analyse their performance often. In order to do this, it may be essential to gather customer input, examine performance metrics, and make any necessary modifications and tweaks to the product in order to improve the overall user experience.

Recommendations for End-Users

It is important for users to offer active feedback about their interactions with chatbot systems. This input may be extremely helpful for helping developers and researchers understand the needs of users, identify problems, and make adjustments that are necessary. Users should exercise caution when disclosing sensitive information when interacting with chatbot systems because these programmes are designed to give experiences that are both useful and personalised. It is essential to have an awareness of the data privacy rules and security measures that are put in place by chatbot systems and to exercise caution when supplying information that is either personally identifiable or secret. Users should explore the potential integration options offered by A.I chatbots with various kinds of platforms and services like mobile apps. By granting access to a greater variety of features and resources, chatbot platforms that provide seamless integration have the potential to boost both worker productivity and user convenience.

REFERENCES

A.Shaji George, A.S.Hovan George, & A.S.Gabrio Martin. (2023). A review of ChatGPT AI's impact on several business sectors. *Partners Universal International Innovation Journal*, *01*(01), 9–23. https://doi.org/10.5281/zenodo.7644359

Abdellatif, A., Badran, K., Costa, D. E., & Shihab, E. (2022). A comparison of natural language understanding platforms for chatbots in software engineering. *IEEE Transactions on Software Engineering*, *48*(8), 3087–3102. https://doi.org/10.1109/TSE.2021.3078384

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

Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. In *IFIP Advances in Information and Communication Technology* (pp. 373–383). https://doi.org/10.1007/978-3-030-49186-4_31

Adamopoulou, E., & Moussiades, L. (2020b). Chatbots: History, technology, and applications. *Machine Learning with Applications*, *2*, 100006. https://doi.org/10.1016/j.mlwa.2020.100006

Amisha, Malik, P., Pathania, M., & Rathaur, V. (2019). Overview of artificial intelligence in medicine. *Journal of Family Medicine and Primary Care*, *8*(7), 2328. https://doi.org/10.4103/jfmpc.jfmpc_440_19

Andrade, I. M. D., & Tumelero, C. (2022). Increasing customer service efficiency through artificial intelligence chatbot. *Revista de Gestão*, *29*(3), 238–251. https://doi.org/10.1108/REGE-07-2021-0120

Anki, P., Bustamam, A., & Buyung, R. A. (2021). Comparative analysis of performance between multimodal implementation of chatbot based on news classification data using categories. *Electronics*, *10*(21), 2696. https://doi.org/10.3390/electronics10212696

Balas, M., & Ing, E. B. (2023). Conversational AI models for ophthalmic diagnosis: Comparison of ChatGPT and the Isabel pro differential diagnosis generator. *JFO Open Ophthalmology*, *1*, 100005. https://doi.org/10.1016/j.jfop.2023.100005

Bankins, S., & Formosa, P. (2023). The ethical implications of artificial intelligence (AI) for meaningful work, *Journal of Business Ethics*. https://doi.org/10.1007/s10551-023-05339-7

Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as ai conversational partners in language learning. *Applied Sciences*, *12*(17), 8427. https://doi.org/10.3390/app12178427

Bhardwaz, S., & Kumar, J. (2023). An extensive comparative analysis of chatbot technologies—chatGPT, Google bard and Microsoft Bing. *In Proceedings of the 2nd International Conference on Applied Artificial Intelligence and Computing*, (pp. 673–679). https://doi.org/10.1109/ICAAIC56838.2023.10140214

Caldarini, G., Jaf, S., & McGarry, K. (2022). A literature survey of recent advances in chatbots. *Information*, *13*(1), 41. https://doi.org/10.3390/info13010041

Camilleri, M. A., & Troise, C. (2023). Live support by chatbots with artificial intelligence: A future research agenda. *Service Business*, *17*(1), 61–80. https://doi.org/10.1007/s11628-022-00513-9

Casas, J., Tricot, M.-O., Abou Khaled, O., Mugellini, E., & Cudré-Mauroux, P. (2020). Trends & methods in chatbot evaluation. *In Proceedings of the Companion Publication of the 2020 International Conference on Multimodal Interaction*, (pp. 280–286). https://doi.org/10.1145/3395035.3425319

Chakrabortty, R. K., Abdel-Basset, M., & Ali, A. M. (2023). A multi-criteria decision analysis model for selecting an optimum customer service chatbot under uncertainty. *Decision Analytics Journal*, *6*, 100168. https://doi.org/10.1016/j.dajour.2023.100168

Chellappa, R. (2022). Can we trust AI?. Baltimore, MD. Johns Hopkins University Press.

Cordero, J., Barba-Guaman, L., & Guamán, F. (2022). Use of chatbots for customer service in MSMEs. *Applied Computing and Informatics*. https://doi.org/10.1108/ACI-06-2022-0148

Cordeschi, R. (2007). AI turns fifty: revisiting its origins. *Applied Artificial Intelligence*, *21*(4–5), 259–279. https://doi.org/10.1080/08839510701252304

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (Fifth edition). Los Angeles, LA. SAGE.

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, *6*(2), 94–98. https://doi.org/10.7861/futurehosp.6-2-94

Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2022). Artificial intelligence and business value: A literature review. *Information Systems Frontiers*, *24*(5), 1709–1734. https://doi.org/10.1007/s10796-021-10186-w

Esser, F., & Vliegenthart, R. (2017). Comparative research methods. *The International Encyclopaedia of Communication Research Methods* (pp. 1–22). Wiley. https://doi.org/10.1002/9781118901731.iecrm0035

Gao, Z., & Jiang, J. (2021). Evaluating human-AI hybrid conversational systems with chatbot message suggestions. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 534–544. https://doi.org/10.1145/3459637.3482340

Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A twenty-year data-driven historical analysis. *International Journal of Innovation Studies*, *4*(4), 134–147. https://doi.org/10.1016/j.ijis.2020.09.001

Habehh, H., & Gohel, S. (2021). Machine learning in healthcare. *Current Genomics*, *22*(4), 291–300. https://doi.org/10.2174/1389202922666210705124359

Hacıoğlu, E. (2021). A comparative study on iterative algorithms of almost contractions in the context of convergence, stability and data dependency. *Computational and Applied Mathematics*, *40*(8), 282. https://doi.org/10.1007/s40314-021-01671-8

Haleem, A., Javaid, M., & Singh, R. P. (2022a). An era of ChatGPT as a significant futuristic support tool: a study on features, abilities, and challenges. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, *2*(4), 100089. https://doi.org/10.1016/j.tbench.2023.100089

Haleem, A., Javaid, M., Asim Qadri, M., Pratap Singh, R., & Suman, R. (2022b). Artificial intelligence (AI) applications for marketing: A literature-based study. *International Journal of Intelligent Networks*, *3*, 119–132. https://doi.org/10.1016/j.ijin.2022.08.005

Han, X., Zhang, Z., Ding, N., Gu, Y., Liu, X., Huo, Y., Qiu, J., Yao, Y., Zhang, A., Zhang, L., Han, W., Huang, M., Jin, Q., Lan, Y., Liu, Y., Liu, Z., Lu, Z., Qiu, X.,

Song, R., … Zhu, J. (2021). Pre-trained models: Past, present and future. *AI Open, 2*, 225–250. https://doi.org/10.1016/j.aiopen.2021.08.002

Hasal, M., Nowaková, J., Ahmed Saghair, K., Abdulla, H., Snášel, V., & Ogiela, L. (2021). Chatbots: Security, privacy, data protection, and social aspects. *Concurrency and Computation: Practice and Experience*, *33*(19). https://doi.org/10.1002/cpe.6426

Haugeland, I. K. F., Følstad, A., Taylor, C., & Bjørkli, C. A. (2022). Understanding the user experience of customer service chatbots: an experimental study of chatbot interaction design. *International Journal of Human-Computer Studies, 161*, 102788. https://doi.org/10.1016/j.ijhcs.2022.102788

Huawei Technologies Co., Ltd. (2023). A general introduction to artificial intelligence. In: *Artificial Intelligence Technology. Springer, Singapore*. https://doi.org/10.1007/978-981-19-2879-6_1

Jadeja, M., & Varia, N. (2017). Perspectives for Evaluating Conversational AI. *ArXiv*. https://doi.org/10.48550/ARXIV.1709.04734

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications*, *82*(3), 3713–3744. https://doi.org/10.1007/s11042-022-13428-4

King, M. R. & chatGPT. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and Molecular Bioengineering*, *16*(1), 1–2. https://doi.org/10.1007/s12195-022-00754-8

Kooli, C. (2023). Chatbots in education and research: a critical examination of ethical implications and solutions. *Sustainability*, *15*(7), 5614. https://doi.org/10.3390/su15075614

Kumar, J. A. (2021). Educational chatbots for project-based learning: Investigating learning outcomes for a team-based design course. *International Journal of Educational Technology in Higher Education*, *18*(1), 65. https://doi.org/10.1186/s41239-021-00302-w

Lin, C.-C., Huang, A. Y. Q., & Yang, S. J. H. (2023). A review of ai-driven conversational chatbots implementation methodologies and challenges (1999–2022). *Sustainability*, *15*(5), 4012. https://doi.org/10.3390/su15054012

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable ai: a review of machine learning interpretability methods. *Entropy*, *23*(1), 18. https://doi.org/10.3390/e23010018

Misischia, C. V., Poecze, F., & Strauss, C. (2022). Chatbots in customer service: their relevance and impact on service quality. *Procedia Computer Science*, *201*, 421–428. https://doi.org/10.1016/j.procs.2022.03.055

Mittal, A., Agrawal, A., Chouksey, A., Shriwas, R., & Agrawal, S. (2016). A comparative study of chatbots and humans. *International Journal of Advanced Research in Computer and Communication Engineering 5*(3). https://doi.org/10.17148/IJARCCE.2016.53253

Mukhamediev, R. I., Popova, Y., Kuchin, Y., Zaitseva, E., Kalimoldayev, A., Symagulov, A., Levashenko, V., Abdoldina, F., Gopejenko, V., Yakunin, K., Muhamedijeva, E., & Yelis, M. (2022). Review of artificial intelligence and machine learning technologies: Classification, restrictions, opportunities and challenges. *Mathematics*, *10*(15), 2552. https://doi.org/10.3390/math10152552

Naik, N., Hameed, B. M. Z., Shetty, D. K., Swain, D., Shah, M., Paul, R., Aggarwal, K., Ibrahim, S., Patil, V., Smriti, K., Shetty, S., Rai, B. P., Chlosta, P., & Somani, B. K. (2022). Legal and ethical consideration in artificial intelligence in healthcare: Who takes responsibility? *Frontiers in Surgery*, *9*, 862322. https://doi.org/10.3389/fsurg.2022.862322

Nicolescu, L., & Tudorache, M. T. (2022). Human-computer interaction in customer service: The experience with AI chatbots—A systematic literature review. *Electronics*, *11*(10), 1579. https://doi.org/10.3390/electronics11101579

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, *2*, 100033. https://doi.org/10.1016/j.caeai.2021.100033

Olujimi, P. A., & Ade-Ibijola, A. (2023). NLP techniques for automating responses to customer queries: A systematic review. *Discover Artificial Intelligence*, *3*(1), 20. https://doi.org/10.1007/s44163-023-00065-5

Pandey, S., & Sharma, S. (2023). A comparative study of retrieval-based and generative-based chatbots using deep learning and machine learning. *Healthcare Analytics*, 100198. https://doi.org/10.1016/j.health.2023.100198

Patil, A., Marimuthu, Rao, N., & Niranchana. (2017). Comparative study of cloud platforms to develop a chatbot. *International Journal of Engineering & Technology*, *6*(3), 57. https://doi.org/10.14419/ijet.v6i3.7628

Perifanis, N.-A., & Kitsios, F. (2023). Investigating the influence of artificial intelligence on business value in the digital era of strategy: A literature review. *Information*, *14*(2), 85. https://doi.org/10.3390/info14020085

Peyton, K., & Unnikrishnan, S. (2023). A comparison of chatbot platforms with the state-of-the-art sentence BERT for answering online student FAQs. *Results in Engineering*, *17*, 100856. https://doi.org/10.1016/j.rineng.2022.100856

Pflanzer, M., Traylor, Z., Lyons, J. B., Dubljević, V., & Nam, C. S. (2022). Ethics in human–AI teaming: Principles and perspectives. *AI and Ethics*. https://doi.org/10.1007/s43681-022-00214-z

Pugliese, R., Regondi, S., & Marini, R. (2021). Machine learning-based approach: Global trends, research directions, and regulatory standpoints. *Data Science and Management*, *4*, 19–29. https://doi.org/10.1016/j.dsm.2021.12.002

Ritchie, J. (Ed.). (2014). Qualitative research practice: A guide for social science students and researchers (2. ed). Sage.

Rodrigues, J., & Lopes, C. T. (2022). Solutions for data sharing and storage: A comparative analysis of data repositories. *Linking Theory and Practice of Digital Libraries* (pp. 512–517). https://doi.org/10.1007/978-3-031-16802-4_55

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, *2*(3), 160. https://doi.org/10.1007/s42979-021-00592-x

Sarker, I. H. (2022). Ai-based modelling: Techniques, applications and research issues towards automation, intelligent and smart systems. *SN Computer Science*, *3*(2), 158. https://doi.org/10.1007/s42979-022-01043-x

Sebastian, G. (2023). Privacy and data protection in ChatGPT and other AI chatbots: Strategies for securing user information. *Social Science Research Network Electronic Journal*. https://doi.org/10.2139/ssrn.4454761

Sharma, M., Verma, S., & Sahni, L. (2020). Comparative analysis of chatbots. *Social Science Research Network Electronic Journal*. https://doi.org/10.2139/ssrn.3563674

Sheikh, H., Prins, C., & Schrijvers, E. (2023). Artificial intelligence: Definition and background. *Mission AI.* (pp. 15–41). https://doi.org/10.1007/978-3-031-21448-6_2

Suta, P., Lan, X., Wu, B., Mongkolnam, P., & Chan, J. H. (2020). An overview of machine learning in chatbots. *International Journal of Mechanical Engineering and Robotics Research*, 502–510. https://doi.org/10.18178/ijmerr.9.4.502-510

Vijayaraghavan, V., Cooper, J. B., & Leevinson, R. (2020). Algorithm inspection for chatbot performance evaluation. *Procedia Computer Science*, *171*, 2267–2274. https://doi.org/10.1016/j.procs.2020.04.245

Vishwakarma, A. (2021). A review & comparative analysis on various chatbots design. *International Journal of Computer Science and Mobile Computing*, *10*(2), 72–78. https://doi.org/10.47760/ijcsmc.2021.v10i02.011

Wang, H., Li, J., Wu, H., Hovy, E., & Sun, Y. (2022). Pre-trained language models and their applications. *Engineering*. https://doi.org/10.1016/j.eng.2022.04.024

Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q.-L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *Journal of Automatica Sinica*, *10*(5), 1122–1136. https://doi.org/10.1109/JAS.2023.123618

Xu, L., Sanders, L., Li, K., & Chow, J. C. L. (2021a). Chatbot for health care and oncology applications using artificial intelligence and machine learning: Systematic review. *JMIR Cancer*, *7*(4), e27850. https://doi.org/10.2196/27850

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., Qiu, C.-W., Qiu, J., Hua, K., Su, W., Wu, J., Xu, H., Han, Y., Fu, C., Yin, Z., Liu, M., … Zhang, J. (2021b). Artificial intelligence: A powerful paradigm for scientific

research. *The Innovation*, *2*(4), 100179. https://doi.org/10.1016/j.xinn.2021.100179

Zhang, J., Oh, Y. J., Lange, P., Yu, Z., & Fukuoka, Y. (2020). Artificial intelligence chatbot behaviour change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet: Viewpoint. *Journal of Medical Internet Research*, *22*(9), e22845. https://doi.org/10.2196/22845

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., … Wen, J.-R. (2023). A survey of large language models. *ArXiv*. http://arxiv.org/abs/2303.18223

Zhou, C., Li, Q., Li, C., Yu, J., Liu, Y., Wang, G., Zhang, K., Ji, C., Yan, Q., He, L., Peng, H., Li, J., Wu, J., Liu, Z., Xie, P., Xiong, C., Pei, J., Yu, P. S., & Sun, L. (2023). A comprehensive survey on pretrained foundation models: a history from BERT to ChatGPT. *ArXiv*. http://arxiv.org/abs/2302.09419

APPENDICES

Appendix A

Ethical Committee Approval

NEAR EAST UNIVERSITY

SCIENTIFIC RESEARCH ETHICS COMMITTEE

20.06.2023

Dear Omar Hussain
Your project **"Comparative Analysis of Artificial Intelligent Chatbots"** has been evaluated. Since only secondary data will be used the project does not need to go through the ethics committee. You can start your research on the condition that you will use only secondary data.

Prof. Dr. Aşkın KİRAZ

The Coordinator of the Scientific Research Ethics Committee

Appendix B

Turnitin Similarity Report

MASTER THESIS

ORIGINALITY REPORT

7%	6%	3%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.researchgate.net Internet Source	1%
2	Saumyamani Bhardwaz, Jitender Kumar. "An Extensive Comparative Analysis of Chatbot Technologies - ChatGPT, Google BARD and Microsoft Bing", 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2023 Publication	1%
3	rgu-repository.worktribe.com Internet Source	<1%
4	vocal.media Internet Source	<1%
5	omniscien.com Internet Source	<1%
6	Submitted to Dundalk Institute of Technology Student Paper	<1%
7	tojqih.net Internet Source	<1%

8	8thinternationalcongressoni.sched.com Internet Source	<1 %
9	Thi Ngoc Anh Tran, Truong Nguyen Ngoc. "Mobile E-Portfolios on Google Sites: A Tool for Enhancing Project-Based Learning", International Journal of Interactive Mobile Technologies (iJIM), 2023 Publication	<1 %
10	Submitted to University of Sunderland Student Paper	<1 %
11	Submitted to Campbellsville University Student Paper	<1 %
12	www.geeksforgeeks.org Internet Source	<1 %
13	rajaic.com Internet Source	<1 %
14	ris.utwente.nl Internet Source	<1 %
15	www.ros.hw.ac.uk Internet Source	<1 %
16	arxiv.org Internet Source	<1 %
17	jurnal.polgan.ac.id Internet Source	<1 %

18	www.ijraset.com Internet Source	<1 %
19	interestingengineering.com Internet Source	<1 %
20	"The Palgrave Handbook of Interactive Marketing", Springer Science and Business Media LLC, 2023 Publication	<1 %
21	Submitted to California State University, Fresno Student Paper	<1 %
22	Saleh Albahli, Aun Irtaza, Tahira Nazir, Awais Mehmood, Ali Alkhalifah, Waleed Albattah. "A Machine Learning Method for Prediction of Stock Market Using Real-Time Twitter Data", Electronics, 2022 Publication	<1 %
23	Sumit Pandey, Srishti Sharma. "A comparative study of retrieval-based and generative-based chatbots using Deep Learning and Machine Learning", Healthcare Analytics, 2023 Publication	<1 %
24	app.clinally.com Internet Source	<1 %
25	Submitted to University of Lancaster Student Paper	<1 %

26	Submitted to University of Ulster Student Paper	<1 %
27	www.artificial-intelligence.blog Internet Source	<1 %
28	www.medrxiv.org Internet Source	<1 %
29	www.oracle.com Internet Source	<1 %
30	www.softwaretestingclass.com Internet Source	<1 %
31	Hossein Hassani, Emmanuel Sirmal Silva. "The Role of ChatGPT in Data Science: How AI-Assisted Conversational Interfaces Are Revolutionizing the Field", Big Data and Cognitive Computing, 2023 Publication	<1 %
32	ijsrp.org Internet Source	<1 %
33	prism.ucalgary.ca Internet Source	<1 %
34	"Chapter 300590 SDG 6", Springer Science and Business Media LLC, 2022 Publication	<1 %
35	crln.acrl.org Internet Source	<1 %

36	dk.um.si Internet Source	<1 %
37	e-journals.unmul.ac.id Internet Source	<1 %
38	ijeltal.org Internet Source	<1 %
39	nou.edu.ng Internet Source	<1 %
40	repositorium.hs-ruhrwest.de Internet Source	<1 %
41	repository.up.ac.za Internet Source	<1 %
42	www.kosen-k.go.jp Internet Source	<1 %
43	www.rsisinternational.org Internet Source	<1 %
44	"Trends and Applications in Knowledge Discovery and Data Mining", Springer Science and Business Media LLC, 2021 Publication	<1 %
45	archive.cm.mahidol.ac.th Internet Source	<1 %
46	community.esri.com Internet Source	<1 %

47	os.unil.cloud.switch.ch Internet Source	<1 %
48	submissao.semead.com.br Internet Source	<1 %
49	transnationalmarket.com Internet Source	<1 %
50	www.grafiati.com Internet Source	<1 %
51	www.tandfonline.com Internet Source	<1 %
52	Blaber, Amanda. "EBOOK: Blaber's Foundations for Paramedic Practice: A Theoretical Perspective", EBOOK: Blaber's Foundations for Paramedic Practice: A Theoretical Perspective, 2018 Publication	<1 %
53	Chengzhong Liu, Shixu Zhou, Yuanhao Zhang, Dingdong Liu, Zhenhui Peng, Xiaojuan Ma. "Exploring the Effects of Self-Mockery to Improve Task-Oriented Chatbot's Social Intelligence", Designing Interactive Systems Conference, 2022 Publication	<1 %
54	Sajid Ali, Tamer Abuhmed, Shaker El-Sappagh, Khan Muhammad et al. "Explainable Artificial Intelligence (XAI): What we know and what is	<1 %

	left to attain Trustworthy Artificial Intelligence", Information Fusion, 2023 Publication	
55	aaltodoc.aalto.fi Internet Source	<1 %
56	assets.researchsquare.com Internet Source	<1 %
57	docs.google.com Internet Source	<1 %
58	edepositireland.ie Internet Source	<1 %
59	hdl.handle.net Internet Source	<1 %
60	icdigital.org.za Internet Source	<1 %
61	qa-jmir.jmir.org Internet Source	<1 %
62	researchbank.swinburne.edu.au Internet Source	<1 %
63	scholarworks.lib.csusb.edu Internet Source	<1 %
64	spectrum.library.concordia.ca Internet Source	<1 %
65	uobrep.openrepository.com Internet Source	<1 %

66	www.africanscholarpublications.com Internet Source	<1 %
67	www.bartleby.com Internet Source	<1 %
68	www.coursehero.com Internet Source	<1 %
69	www.frontiersin.org Internet Source	<1 %
70	www.ijtsrd.com Internet Source	<1 %
71	www.mdpi.com Internet Source	<1 %
72	www.scilit.net Internet Source	<1 %
73	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 %
74	Sofia Schöbel, Anuschka Schmitt, Dennis Benner, Mohammed Saqr, Andreas Janson, Jan Marco Leimeister. "Charting the Evolution and Future of Conversational Agents: A Research Agenda Along Five Waves and New	<1 %

	Frontiers", Information Systems Frontiers, 2023 Publication	
75	Kevin Peyton, Saritha Unnikrishnan. "A comparison of chatbot platforms with the state-of-the-art sentence BERT for answering online student FAQs", Results in Engineering, 2022 Publication	<1 %
76	www.ieee-jas.net Internet Source	<1 %

Exclude quotes	On	Exclude matches	< 5 words
Exclude bibliography	On		