SENTIMENT ANALYSIS TECHNIQUE FOR THE DEVELOPMENT OF NEW IMPROVED CHATBOT: NEW USER HUMANITY-LIKE CHATBOT (NURUL-C)

NURUL MUIZZAH BINTI JOHARI

UNIVERSITI KEBANGSAAN MALAYSIA

SENTIMENT ANALYSIS TECHNIQUE FOR THE DEVELOPMENT OF NEW IMPROVED CHATBOT: NEW USER HUMANITY-LIKE CHATBOT (NURUL-C)

NURUL MUIZZAH BINTI JOHARI

THESIS SUBMITTED IN FULFILMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

INSTITUTE OF VISUAL INFORMATICS
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2025

TEKNIK ANALISIS SENTIMEN UNTUK PEMBANGUNAN BOT SEMBANG BAHARU YANG DIPERTINGKATKAN: *NEW USER HUMANITY-LIKE CHATBOT (NURUL-C)*

NURUL MUIZZAH BINTI JOHARI

TESIS YANG DIKEMUKAKAN UNTUK MEMPEROLEH IJAZAH DOKTOR FALSAFAH

INSTITUT VISUAL INFORMATIK
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2025

UNIVERSITI KEBANGSAAN MALAYSIA

The National University of Malaysia

PERAKUAN TESIS SARJANA / DOKTOR FALSAFAH
(DECLARATION OF MASTER / DOCTOR OF PHILOSOPHY THESIS)

A. MAKLUMAT PELAJAR/*STUDENT DETAILS*	
Nama *Name*	NURUL MUIZZAH BINTI JOHARI
No. Pendaftaran *Registration No.*	P94611
Fakulti/Institut *Faculty/Institute*	INSTITUT VISUAL INFORMATIK
Program Pengajian *Program of Study*	~~Sarjana~~ / **Doktor Falsafah** | ~~*Masters*~~ / ***Doctor of Philosophy***
Tajuk Tesis *Thesis Title*	SENTIMENT ANALYSIS TECHNIQUE FOR THE DEVELOPMENT OF NEW IMPROVED CHATBOT: NEW USER HUMANITY-LIKE CHATBOT (NURUL-C)
Mel-e/*Email*: muizjohari91@gmail.com	No. Telefon/*Tel.No.*: 017-563 7551

B. PERAKUAN / *DECLARATION*

Merujuk kepada Dasar Harta Intelek UKM 2018, **tesis adalah hak milik UKM** seperti keterangan dibawah:

4.2.1 Harta Intelek yang direka cipta oleh pelajar termasuk tesis dan semua hasil penyelidikan ditulis oleh Pelajar UKM dalam tempoh pengajiannya di UKM direka cipta, dibangunkan atau dihasilkan menggunakan kemudahan, bahan, dana, atau sumber lain adalah hak milik UKM melainkan dinyatakan sebaliknya dalam Dasar ini atau dipersetujui sebaliknya oleh UKM dan Pelajar/Penaja.

Referring to the UKM Intellectual Property Policy 2018, ***the thesis is the property of UKM*** *as described below:*

4.2.1 Intellectual property created by the student, including the thesis and all research output produced by the UKM Student during their studies at UKM, including designed, developed or produced output using facilities, materials, funds, or other resources are the property of UKM unless otherwise stated in this Policy or otherwise agreed by UKM and the Student/Sponsor.

	RAHSIA *CONFIDENTIAL*	Mengandungi maklumat rahsia di bawah AKTA RAHSIA RASMI 1972 *Consisting of classified information under the OFFICIAL SECRETS ACT 1972*
	TERHAD *RESTRICTED*	Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi / badan di mana penyelidikan dijalankan *Consisting of RESTRICTED information which has been determined by the organisation/body where the research was conducted*
/	AKSES TERBUKA/ TIDAK TERHAD *OPEN ACCESS/ NON-RESTRICTED*	Saya membenarkan tesis ini diterbitkan secara akses terbuka, teks penuh atau dibuat salinan untuk tujuan pengajian, pembelajaran & penyelidikan sahaja *I give permission for this thesis to be published through open access, full text or copied only for study, learning, and research purposes.*

Bagi kategori Akses Terbuka/Tidak Terhad, saya membenarkan tesis (Sarjana/Doktor Falsafah) ini di simpan di Perpustakaan Universiti Kebangsaan Malaysia (UKM)* dengan syarat-syarat kegunaan seperti berikut: *For the Open Access/Non-Restricted category, I permit this (Master's/Doctoral) Thesis to be kept in the Universiti Kebangsaan Malaysia (UKM) Library with the following conditions:* 1. Perpustakaan UKM mempunyai hak untuk membuat salinan untuk tujuan pengajian, pembelajaran, penyelidikan sahaja. *UKM Library has the right to reproduce the thesis only for study, learning and research purposes.* 2. Perpustakaan Universiti Kebangsaan Malaysia dibenarkan membuat satu (1) salinan tesis ini untuk tujuan pertukaran antara institusi pengajian tinggi dan mana-mana badan/ agensi kerajaan, tertakluk kepada terma dan syarat. *UKM Library is allowed to make one (1) copy of this thesis for exchange purposes among higher education institutions and any government body/agency, subject to the terms and conditions.*

Saya mengakui maklumat & tesis yang diberikan adalah benar & terkini. *I acknowledge the information & the thesis provided are correct & current.* Tandatangan Pelajar: *Student's Signature* *Tarikh / Date: 22/10/2025*	Saya memperakukan maklumat & tesis yang diberikan adalah benar & terkini. *I verify that the information & the thesis provided are correct & current.* Tandatangan Penyelia / *Supervisor's Signature*: Nama / *Name*: Dr. Ahmad Hanif Ahmad Baharin Cop Rasmi / *Official Stamp*: *Tarikh / Date:* 22/10/2025

C. PENGESAHAN FAKULTI/INSTITUT | *VERIFICATION OF FACULTY/INSTITUTE*
Tandatangan/*Signature*: Nama/*Name*: Tarikh/*Date*: 20/11/2025 Cop Rasmi/*Official Stamp*: MOHAMAD SALLEH SULIEMAN Penolong Pendaftar Kanan Institut Informatik Visual (IVI) Universiti Kebangsaan Malaysia

**Kuatkuasa:15 Julai 2021

DECLARATION

I hereby declare that the work in this thesis is my own except for quotations, and summaries which have been duly acknowledged.

20 November 2025

NURUL MUIZZAH BINTI JOHARI
P94611

ACKNOWLEDGMENT

In the name of Allah, the Most Gracious and the Most Merciful.

Alhamdulillah, all praise to the Almighty God for the strength and blessing given to me to complete this PhD journey throughout these challenging yet meaningful years. The support and guidance provided by the people surrounding me have driven me to this level. I would like to express my high appreciation to Dr Ahmad Hanif Ahmad Baharin for his support, patience, and guidance for me to keep moving forward and completing this thesis regardless of the hurdles faced at points of the journey. Other personnel who contributed to this fruitful journey were my co-supervisor, Dr Puteri Nor Ellyza Nohuddin, who currently dedicated her time while in Malaysia as well as in the United Arab Emirates to guide me and provide ample support for me to finish the journey. Lastly, I would also like to express my greatest gratitude to Professor Dato' Dr. Halimah Badioze Zaman who brought the first sparks for me to start the journey in 2018. The appreciation goes to her time, warm support, and dedication to keep me on the right track throughout the journey. All suggestions, motivation, as well as encouragement given by my supervisors, have brought me to this stage these days. It was a privilege and a big honour to work and study under their guidance.

The gratitude also to the members of the Institute of Visual Informatics (IVI) and the Graduate Center of Universiti Kebangsaan Malaysia (UKM) for their help and contribution throughout my years in UKM and for providing pleasant working environments as well as the students of UKM who have taken part in this research.

My sincere thanks also go to my friends and colleagues who have given their support and cooperation along this PhD journey including my valued friends from Universiti Malaysia Sarawak (UNIMAS), Mohd Hafizan Hashim and Rafika Jamaluddin who have inspired me to step forward.

Finally, to my dearest parent, Johari Ismail and Zaidah Zakaria who showered me with their prayers, blessings, support, and never-ending encouragement for me to finish the journey that I have taken. Not to forget my lovely sisters, Nurul Farhin and Nurul Hannani as well as my charming brothers, Muhammad Muizzuddin and Muhammad Amir Syauqi for their companionship and strength to lend to me when I need them the most.

ABSTRACT

Emerging technologies increasingly simplify daily human tasks, with chatbots—powered by Artificial Intelligence (AI) and Natural Language Processing (NLP)—being among the most common. Although designed to simulate human conversation, chatbots often fail to meet user needs due to limitations in training data, leading to misunderstandings. This research addresses four gaps: identifying essential chatbot quality attributes, developing a text analyser, implementing sentiment analysis, and validating user satisfaction with a newly developed chatbot. Grounded in information-seeking behaviour theory and focused on the banking domain, the study followed three phases: pre-development; development, integration, and implementation; and evaluation and testing, using an iterative approach. In the pre-development phase, a pilot study with 32 customers helped refine the key quality attributes. Six were identified—functionality, efficiency, technical satisfaction, humanity, effectiveness, and ethics—with two additional attributes added later: visual elements and energy efficiency. These informed the development of the improved chatbot, NURUL-C. The system incorporated text analysis techniques such as stemming and lemmatization. In final evaluations, NURUL-C received positive feedback, supported by favourable sentiment analysis and high user satisfaction. The inclusion of energy efficiency and visual elements contributed to these outcomes. A major contribution of this work is the Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle (HEI-HC-LC), which offers significant implications for service application development in the banking sector. Two (2) research hypotheses were tested. The first null hypothesis, *"There is no difference between quality attributes of a good chatbot and customer satisfaction levels"*, was rejected because there was a strong and significant correlation between chatbot attributes and customer satisfaction ($r = .7689$, $p = .0093$). The second null hypothesis (H02), *"There is no relationship between positive sentiment analysis and customer satisfaction levels"*, could not be rejected due to a weak and non-significant positive correlation between positive sentiment and customer satisfaction ($r = .274$, $p = .443$). Overall, users were satisfied with the new chatbot, NURUL-C.

TEKNIK ANALISIS SENTIMEN UNTUK PEMBANGUNAN BOT SEMBANG BAHARU YANG DIPERTINGKATKAN: *NEW USER HUMANITY-LIKE CHATBOT (NURUL-C)*

ABSTRAK

Dunia masa kini dipacu oleh kemajuan teknologi yang bertujuan mempermudah urusan harian manusia. Antara teknologi yang mengintegrasikan Kecerdasan Buatan (AI) dan Pemprosesan Bahasa Tabii (NLP) ialah bot sembang (*chatbot*), iaitu aplikasi yang mensimulasikan komunikasi manusia melalui bahasa tabii. Namun, interaksi pengguna–sistem sering menimbulkan salah faham disebabkan kekangan data latihan, sekali gus menghadkan keupayaan bot sembang memenuhi keperluan sebenar pengguna. Penyelidikan ini menangani empat jurang utama: mengenal pasti atribut yang berkualiti bagi bot sembang yang berkesan, membangunkan penganalisis teks, melaksanakan analisis sentimen dan menilai kepuasan pengguna terhadap bot sembang baharu. Berasaskan teori tingkah laku mencari maklumat dalam konteks perbankan, kajian dilaksanakan melalui tiga fasa—pra-pembangunan; pembangunan, integrasi dan pelaksanaan; serta penilaian dan ujian—menggunakan pendekatan iteratif bagi memastikan pencapaian objektif secara menyeluruh. Kajian rintis melibatkan 32 responden berjaya memperincikan atribut utama untuk pembangunan sistem. Enam atribut telah dikenal pasti: kefungsian, kecekapan, kepuasan teknikal, kemanusiaan, keberkesanan dan etika; disusuli dua atribut tambahan sebelum pembangunan bot sembang baharu, iaitu elemen visual serta penjimatan dan kecekapan tenaga. Berdasarkan penemuan ini, NURUL-C dibangunkan dan teks interaksinya dianalisis melalui teknik pelucutan imbuhan dan lematisasi. Hasil penilaian akhir menunjukkan NURUL-C menerima maklum balas positif melalui analisis sentimen dan tahap kepuasan pengguna yang memberangsangkan. Penyepaduan elemen visual dan kecekapan tenaga turut menyumbang kepada peningkatan prestasi sistem. Sumbangan utama kajian ialah penghasilan Kitaran Hayat Pembangunan Bot Sembang Kemanusiaan Berevolusi Holistik (HEI-HC-LC), yang menawarkan implikasi signifikan terhadap pembangunan aplikasi perkhidmatan, khususnya dalam sektor perbankan. Terdapat dua (2) hipotesis kajian yang telah diuji. Pengujian hipotesis nul pertama iaitu “Tiada perbezaan antara atribut berkualiti bagi bot sembang yang baik dan tahap kepuasan pelanggan” mendapati hipotesis tersebut ditolak kerana terdapat korelasi kuat dan signifikan antara atribut bot sembang dan kepuasan pelanggan ($r = .7689$, $p = .0093$). Hipotesis nul kedua (H02) iaitu “Tiada perhubungan antara analisis sentimen positif dan tahap kepuasan pelanggan” tidak dapat ditolak berikutan korelasi positif yang lemah dan tidak signifikan antara sentimen positif dan kepuasan pelanggan ($r = .274$, $p = .443$). Secara keseluruhan, pengguna berpuas hati dengan bot sembang baharu, NURUL-C.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A.L.I.C.E	Artificial Intelligence Internet Computer Entity
ABSA	Aspect-Based Sentiment Analysis
ACD	Aspect Category Detection
AGQC	Attribute Good Quality Chatbot
AI	Artificial Intelligence
AMOS	Analysis of Moment Structure Software
AWS	Amazon Web Services
BoW	Bag-of-Words
CB	Customer Behaviour
ChatGPT	Chat Generative Pre-Trained Transformer
CSAT	Customer Satisfaction Score
CSI	Customer Satisfaction Index
ESG	Environmental Social and Governance
FAQ	Frequently Asked Questions
HAAM-C-1	Holistic Attribute Model for Chatbot: Demographic Information
HAAM-C-2	Holistic Attribute Model for Chatbot (Chatbot Evaluation on User's Perception on Effectiveness of Chatbot)
HAAM-C-3	Holistic Attribute Model for Chatbot (Customer Satisfaction Level on Chatbot Evaluation)
HAAM-C-4	Holistic Attribute Model for Chatbot (Development of Text Analyzer Techniques)

HAAM-C-5	Holistic Attribute Model for Chatbot (Sentiment Analysis)
HAAM-NC	Holistic Attribute Analysis Model for New Chatbot
HAAM-NC-1	Holistic Attribute Analysis Model of New Chatbot: Demographic Information
HAAM-NC-E	Holistic Attribute Analysis Model of New Chatbot Evaluation
HAAM-NC-S1-S3	Holistic Attribute Analysis Model of New Chatbot: Scenario 1- 3
HCI	Human Computer Interaction
HEI-HC-LC	Holistic Evolutionary Iterative Humanistic Chatbot Development Live Cycle
HIB	Human Information Behaviour
IB	Information Behaviour
IDF	Inverse Document Frequency
IoT	Internet of Things
ISB	Information Seeking Behaviour
ML	Machine Learning
MS-DOS	Microsoft Disk Operating System
NER	Named Entity Recognition
NLP	Natural Language Processing
NLU	Natural Language Understanding
NURUL-C	New User Humanity-like Chatbot
OTE	Opinion Target Extraction

QoS	Quality of Service
SP	Sentiment Polarity
SPSS	Statistical Package for Social Science
SRS	System Requirement Specifications
SUS	System Usability Scale
TF	Term Frequency
TF-IDF	Term Frequency – Inverse Document Frequency
UAT	User Acceptance Test
UCD	User Centered Design
UKM	Universiti Kebangsaan Malaysia
UX	User Experience
VCI	Visual Conversation Interface

CHAPTER I

INTRODUCTION

1.1 INTRODUCTION

This chapter explains the research background, problem statement, research questions, objectives, hypotheses, research scope, expected outcomes as well as the significance of the study. There are eight sub-sections in this chapter. Section 1.2 discusses the research background which briefly explains aspects related to chatbot, sentiment analysis as well as customer satisfaction level which eventually relates to Customer Satisfaction Index. These three elements of interest are the keywords that will be discussed throughout this research. Section 1.3 discusses the problem statement which fundamentally indicates the gap found in the 'state-of-the-art' research on chatbot development and usage. Section 1.4 describes the research objectives that became the main reference to answer the research questions developed. Section 1.5 describes the research questions, which analyse further and support the outcomes based on research objectives. Section 1.6 discusses the research theoretical framework while Section 1.7 focuses on the research conceptual framework. Section 1.8 discusses the research scope and Section 1.9 briefs on the significance of the research and its contributions towards the field of knowledge. Section 1.10 addressed the limitation of the research. This chapter will be concluded in Section 1.11 on the thesis organization and Section 1.12 on the chapter conclusion.

1.2 RESEARCH BACKGROUND

Computer-based chatbots are currently getting wider attention in various fields in domains such as banking and finance, entertainment, as well as education. A chatbot is an application that imitates human intelligence (through learning and training) and can reproduce reasonably ‘intelligent’ conversations using natural language (Ranoliya et al. 2017). Chatbots are commonly used for customer service across various industries (Xu et al. 2017). However, the conversation applied to many chatbots is mostly through instant messaging whereby it is accessible to all types of users and not specific to certain types of customers. Since they are developed usually using a user-friendly interface, users are familiar with and take less time to learn how to use the system.

Zaboj’s (2020) finding showed that end users have developed their preferences for chatbots which are expected to assist them in solving minor issues. There are statistics related to chatbots that compared to 2018, the year 2019 indicated the number of end-users who were also consumers doubled in terms of willingness to interact with chatbots as they perceived the application to be immensely helpful to their needs (Press 2019). Seventy-four per cent (74%) of users preferred to use chatbots to search for answers to solve inquiries (PSFK 2021). Sixty-five per cent (65%) of consumers felt at ease in handling issues without the presence of human agents (Bazillian 2017). It was also found through his study that sixty-nine per cent (69%) of consumers chose to use chatbots as they wanted quick feedback based on simple straightforward questions (Sweezey 2019). Interestingly, forty-eight per cent (48%) of the consumers preferred to be engaged with a chatbot that can solve issues than with a chatbot that was designed and developed with personality (Beaver 2017). Interestingly too, sixty-four per cent (64%) of the consumers claimed that 24/7 services were the most useful functionality of a bot so that these chatbots would be available whenever they needed them. (Knight 2019). Eighty-two per cent (82%) of the consumers perceived that instant feedback to any inquiry is of utmost importance when they need answers or solutions quickly (Moyle 2019).

Energy consumption in terms of time taken to complete any searching tasks imposed significant roles in information retrieval. This activity closely relates to the growth in demand for data centres to accommodate all data needed in the information retrieval process through the Internet. Rozite et al. (2023) explain activities such as streaming, cloud gaming, blockchain, artificial intelligence, machine learning, and virtual reality contribute to the escalation of demand for data services. Artificial Intelligence (AI) and machine learning (ML) require a significant amount of computing power to train the models. Based on data from Meta and Google (Patterson et. al. 2022), the training phase consumes twenty (20) to forty per cent (40%) of overall ML-related energy use, with sixty per cent (60%) to seventy per cent (70%) for inference and more than ten per cent (10%) for model development (Rozite et al. 2023). This leads AI to assist in reducing energy use in data centres to cater for the computing demand.

Sentiment is an opinion expressed by someone upon something that matters or is of concern (Delacroix 2007). Sentiment analysis is used to indicate the attitude or perception of users, speakers, or writers based on the topic, content, or overall context of a document (Kowcika et al. 2013). The basic task of sentiment analysis is to classify the direction of the statement given in any text in the form of sentences or phrases as well as emojis whether it points to positive, negative, or neutral perception (Kowcika et al. 2013). The application of sentiment analysis currently is not limited to predicting election results (Tumasjan et al. 2010), but it can also assist in classifying issues within text conversations that can indicate opinions such as customer satisfaction (Xiang et al. 2014).

The need to search for information can be conducted using services provided by a chatbot. The act of searching for information is known as Information Behaviour (IB). This information-searching behaviour is important in human's daily lives which helps solve their problems. Wilson (2000) describes IB as the collective behaviour of humans in connection to sources of information that include both active and passive information seeking and use. Information-seeking approaches are applied in library and information science research (Wilson 2000), and these approaches can be applied to chatbots. Whenever a time limit is applied, it is found that less time consumed for information-seeking is highly able to affect less result check-ups (Chiravirakul & Payne 2014; Crenscenzi et al. 2015), better time for document inspection (Crescenzi et al. 2015), and less information inquired (Liu et al. 2014). Chatbot serves users intending to help users, by providing inquiry information to help solve their problems.

Customer satisfaction levels vary according to customers' experience of using any services provided by an organization. The level can be rated from 'very satisfied' to 'very unsatisfied.' This rating conducted by customers could affect the overall performance of the organisation generally and the chatbot specifically, especially if customers perceived it as easy or difficult to use. The measurements of customer satisfaction vary depending on the goals that need to be achieved. Various measurement metrics can be used to measure customer satisfaction, and the specific one is the Customer Satisfaction (CSAT) metric to measure user's or customer's feelings regarding the functionality and performance of a new chatbot.

1.3 PROBLEM STATEMENT

The interaction between users or customers and the chatbot can be misunderstood as the machine learning ability of a chatbot is limited to the context of the word or phrases related to the main function of the chatbot. For example, a chatbot for banking should identify words or phrases related to withdrawal, to possess a positive meaning as an act of taking out money from 'one's own' bank account. According to Coniam (2014), most existing chatbots fail to meet user's or customers' needs and expectations due to unclear objectives, illogical responses, or lacking effective usability engineering. Yan et al. (2016) stated in their research on serverless chatbots that there were challenges faced pertaining to aspects such as programming and debugging, performance, monitoring, and security. Programming for a serverless Chatbot would take a series of collective functions to create a flow for completing certain tasks. However, there is a lack of existing tools that allow developers to easily debug the program. With technological advancement, serverless platforms such as AWS Lambda allow functions to communicate among themselves over a shared messaging channel.

Another challenge is performance whereby a serverless chatbot makes it hard to provide Quality of Service (QoS) guarantees. To provide QoS guarantees, the serverless platform must be able to communicate the needed QoS with outside components. Serverless Chatbot requires developers to monitor the performance as the deployment of the solution is done in real-time without delay. Lastly, the other challenge is security issues where serverless Chatbot is exposed to cyber-attack and leakage of data and information. A survey recorded by the Chatbot Report 2019 (2019), indicate findings that investigate the problem faced by organisations that use intelligent assistants or chatbots in their workplace. Based on the data collected, it was found that most participants indicated that the chatbots often misunderstand their requests based on the different patterns of human dialogues. Many other misunderstandings have happened with chatbots such as misunderstanding the different distinctions that take place in human dialogue; misunderstanding the requests by customers; giving incorrect command execution; they are unable to understand the different accents of the customers and unable to differentiate the customers' voice as well as the tone of voice that would denote a different meaning. What is also a great challenge is the fact that the chatbots often give incorrect information; the use of incorrect or inappropriate language during feedback and they have been known to release false emergency alarms that had confused customers.

Mimoun et al. (2012) and Oliver (2014) also shared that there are existing chatbots that are unable to meet users' or customers' needs which leads to unsatisfactory customer experience. These downside events have caused negative impacts on organisations generally and service providers specifically in terms of trust, loyalty, as well as intention to purchase the product or service offered. Shah et al. (2017) discusses the issues that will be faced during the development of banking assistant which is the chatbot. During designing, there might be large computational operating cost as it involves vast memories, rapid processors, and in some cases, requiring parallel architectures. There is also an issue to choose the best optimum interface whether to be voice-based or text-based or both. Another concern addressed is that it is time consuming process as it involves designing, collecting data (manually) and testing.

The latest challenge is how developers can design chatbots to be more visual, which in many ways can also address the misunderstanding issues; as well as meet the needs of the world concerns in aspects of Environmental Social and Governance (ESG) as well as sustainability: beginning by incorporating energy efficiency and energy savings in the development of chatbot.

Thus, to overcome all these problems, the plausible solution is the design and development of a new chatbot based on a novel chatbot framework based on novel quality attributes.

1.4 RESEARCH OBJECTIVES

The main aim of this research is to explore the sentiment analysis technique and the satisfaction level of customers on a new improved chatbot, to assess the Customer Satisfaction Index. To achieve these four (4) objectives were developed as follows:

R01 To determine quality attributes for a new chatbot and then validate the chatbot.

R02 To develop text analyser using stemming and lemmatization techniques.

R03 To explore and develop the sentiment analysis based on text analysis available in the chatbot.

R04 To validate the user's or customer's level of satisfaction with the chatbot's text analyser.

1.5 RESEARCH QUESTIONS

The research questions (RQ1 – RQ4) were designed to drive the research direction and to meet the objectives of the research.

i. The first research question (RQ1) will be answered using two approaches, past works based on literature review, and a pilot survey conducted. RQ1 is shown below:

RQ1: What are the attributes of a good quality chatbot?

ii. The second research question (RQ2) focused on the approaches used in designing the new chatbot. The research question is stated as below:

RQ2: What are the approaches used to design and develop the new chatbot?

iii. The third research question (RQ3) focused on sentiment analysis. The data will be extracted from the chatbot conversations and analysed at the back end of the system. The research question is as stated below:

RQ3: What is the sentiment shown by users from the text analysis of the chatbot?

iv. The fourth research question (RQ4) focused on the level of customer satisfaction. Thus concludes the final research objective. The research question is as indicated below:

RQ4: What is the level of satisfaction of the user or customer towards the chatbot?

1.6 RESEARCH HYPOTHESES

Research hypotheses were designed to enhance the findings of the study based on research objectives (R01-R04) and research questions (RQ1-RQ4). The hypotheses that were developed were as follows:

H1 There is no significant difference between quality features for a good chatbot and Customer Satisfaction level.

H2 There is no correlation between positive sentiment analysis and Customer Satisfaction level.

1.7 RESEARCH FRAMEWORK

The motivation of this research is based on information seeking behavior by Wilson (2000) which correlated with Human Information Behavior (HIB) by Krikelas (1958). It is mentioned there are four (4) steps involved in searching information which are identify the needs, searching, finding information, and utilize the information.

In this research, as shown in Figure 1.1 HIB took effect on the research framework that consist of elements such as chatbot, text analysis, sentiment analysis, and user satisfaction. Attributes that become the center of the research focus are functionality, efficiency, technical satisfaction, humanity (co-exist with visual), effectiveness (co-exist with energy efficiency), and ethics. The details are further discussed in Section 1.8 (Research Theoretical Framework) and Section 1.9 (Research Conceptual Framework).

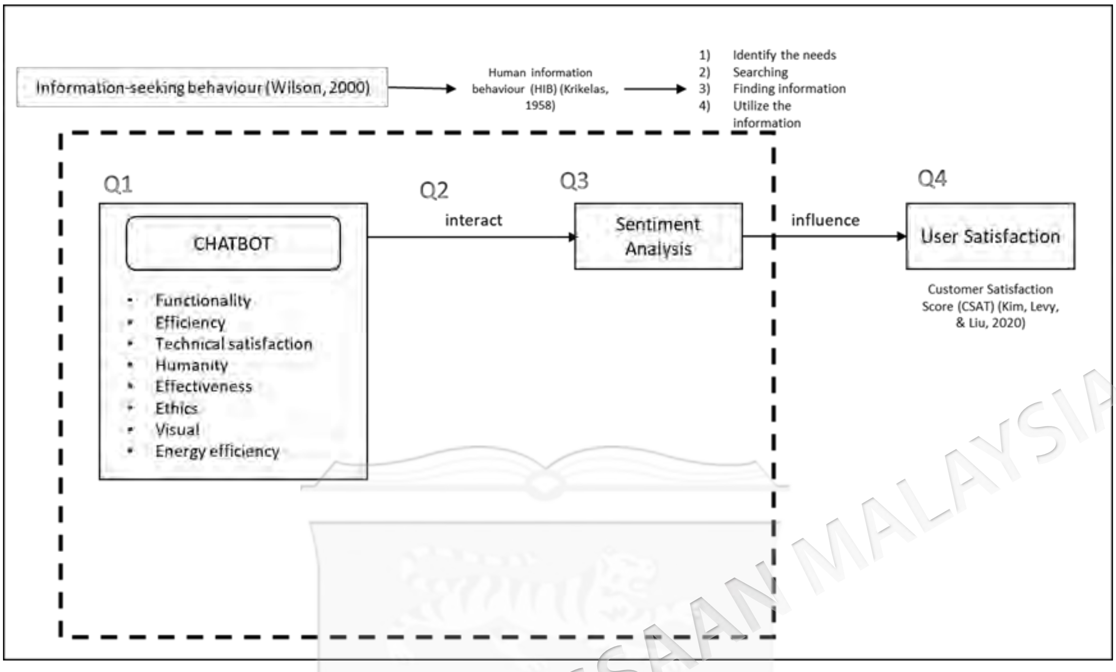


Figure 1.1 Research framework

1.8 RESEARCH THEORETICAL FRAMEWORK

The theoretical framework of the research is as indicated in Figure 1.2 which details the theories referred throughout the research.

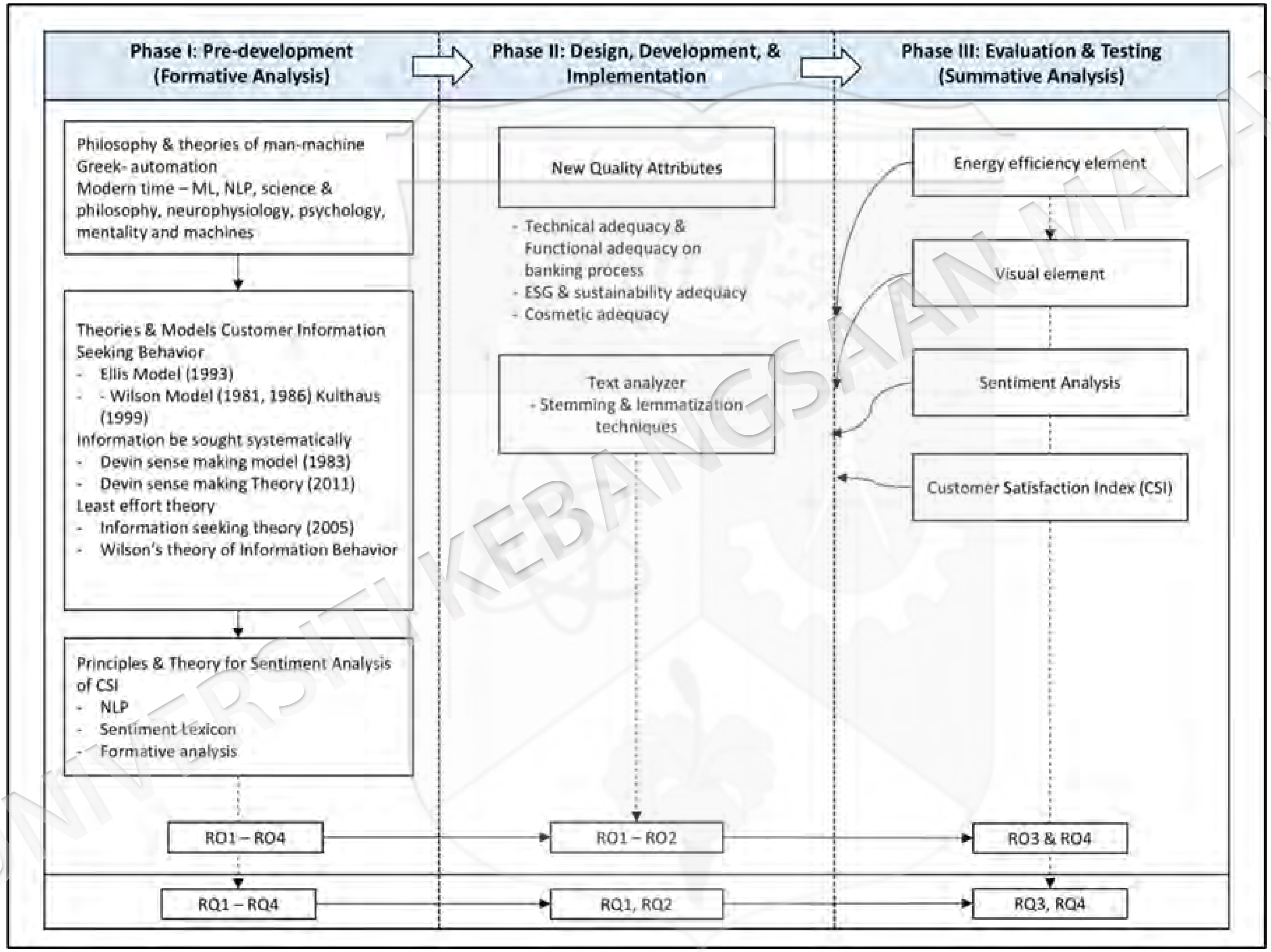


Figure 1.2 Theoretical framework of research

The theoretical framework of this research can be divided into three (3) phases. Firstly, the pre-development phase focused at the first dimension which studied all the possible science and philosophy that is related to human-machine from the time of Greek to the present modern era. It is important to understand the philosophy, science, and related theories to ensure that the development of the chatbot does not only take into consideration the technology aspect but also the humanisation aspect of the design (interface, usage of colour, and others) and development of the chatbot; being constantly conscious that it is a multidisciplinary approach that takes into consideration the technology domain (automation, machine learning, natural language processing, and the like) as well as neurophysiology, psychology, and mentality and machines. Not all were integrated directly into the chatbot but some of the humanistic aspects such as the greetings, tone of voice was consciously aware that it should be taken into consideration by the future developers.

This phase too, studied closely the second dimension that involved the information seeking behaviour theories that was crucial to the design and development of the new chatbot. The content analysis of the theories related to the information seeking behaviour models such as the Ellis Model (1993), Wilson Model (1981, 1986, & 1999) and the Kuhlthau's (1992) model were analysed to see how information can be sought systematically. The third dimension that was incorporated into this phase of the framework was the theory and principles on sentiment analysis and Customer Satisfaction Index (CSI). Sentiment is believed can be a powerful tool to make sense of the vast sea of thought, likes, dislikes, and opinions, which are more than whispers in the wind of more than 5.18 billion humans that are connected via the internet. Out of this, more than 5 billion are engaged in social media. Thus, this dimension is important to consider in the design and development of a new good quality chatbot. All the three (3) dimensions and their elements helped to answer RQ1 through RQ4 of the research.

The second phase of the theoretical framework involved two (2) dimensions namely the new quality attributes and text analyser and lemmatization techniques. The former dimension involved elements such as technical adequacy, functional banking adequacy, the cosmetic adequacy, and the ESG and sustainability adequacy. The latter dimension involved the text analyse and the lemmatisation techniques. Both these dimensions and their specific elements are important to ensure that the design of the chatbot meet the needs of the users or customers of the banking sector. All the two (2) dimensions and their respective elements helped to answer RQ2 and RQ3 of the research.

The third phase of the theoretical framework is the Assessment dimension which encompassed four (4) elements such as energy efficiency element, visual element, sentiment analysis, and customer satisfaction index. The energy efficient element involved the design and assessment of how the chatbot can ensure the use of efficient energy to meet the ESG and sustainability requirement of the organisation. The visual element is to ensure that the design of the chatbot can help users or customers to understand feedback and statements made by the chatbots in a more effective and meaningful manner to avoid any misunderstandings. The sentiment analysis element ensures that measurement on the opinions of the users or customers regarding the design development and the services that the chatbot is supposed to render is positively and effectively portrayed. The sentiment satisfaction index element would be able to convey the levels of satisfaction of the users or customers using the new quality chatbot. All the elements in the assessment dimension helped to answer the RQ3 and RQ4 of the research.

1.9 RESEARCH CONCEPTUAL FRAMEWORK

Like the theoretical framework, the conceptual framework of this research is also divided into three (3) phases which encompass the process to conduct the design and development of the chatbot based on the objectives developed. Details of the phases are described below:

The first phase is the pre-development phase where all information related to the chatbot is collected to identify and determine the quality attributes to be integrated into the design of the new chatbot. The activities involved in this phase were aimed at achieving the first research objective (R01). Two approaches were undertaken to collect the needed information which were namely conducting content analysis through literature review and conducting preliminary survey to analyses data to determine the best quality attributes to be included in the new chatbot.

The second phase involved the development of a chatbot. The activities involved in this phase included creating a library to train the chatbot and incorporating a text analyser and sentiment analysis features in the chatbot. The testing involved in this phase was divided into two tests: pre-development (formative test) and post-development (summative test). This phase was aimed to answer the second and third research objectives (RO2 and RO3). This phase is supported by the information-seeking behaviour and information-seeking theories of the theoretical framework.

The third phase was the analysis or evaluation phase. The process of analysis included the process of investigating the sentiment of the users or customers based on the conversation script collected from the back end of the chatbot as well as measuring the customer experience in terms of their satisfaction based on the service provided by the chatbot.

Figure 1.3 shows the details of the research conceptual framework including the phases and activities conducted to design and develop the new chatbot.

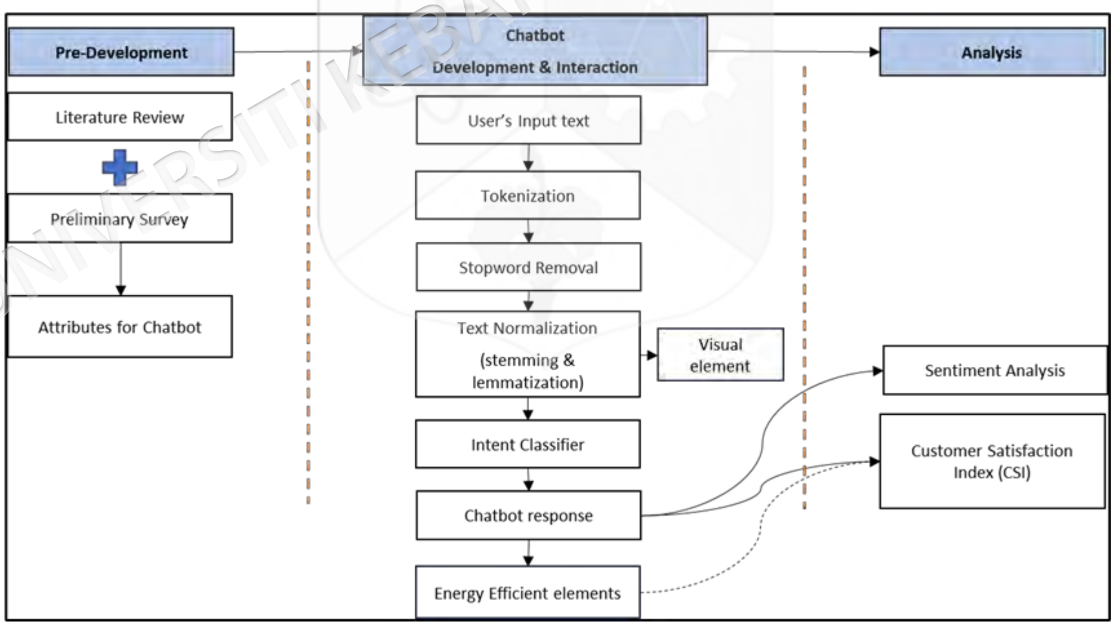


Figure 1.3 Research conceptual framework

1.10 RESEARCH SCOPE

The research conducted focused on the design and development of a new chatbot based on the quality attributes ascertained. The research also focuses on sentiment analysis and customer satisfaction levels through the usage of the new developed chatbot. The chatbot was designed and developed based on the banking domain.

The participants were selected randomly from diverse demographic backgrounds ranged from students to public and private employees as well as self-employed. This research was conducted through several phases which involved conducting a literature review, preliminary survey, and pre-development and post-development (which were conducted after the development of the new chatbot). The evaluation was conducted in the form of a survey. The instruments used for the survey were adapted from past literature and tested for reliability and validity of the items in the instruments to ensure accuracy of the data.

At the end of the research, the outcomes were in the form of sentiment analysis and scores (metrics) for the Customer Satisfaction Index (CSI) used to answer the research objectives (RO1-RO4) as well as the research questions (RQ1-RQ4) and research hypotheses (H1-H2). Table 1.1 describes the summary of the research focused on the research objectives, research questions, and research hypotheses. Figure 1.4 amplifies the research scope which involved the banking domain only; the quality attributes that were identified and validated; the information-seeking behaviour, sentiment analysis and its effect on the Customer Satisfaction Index (CSI). Table 1.1 summarises the research focus.

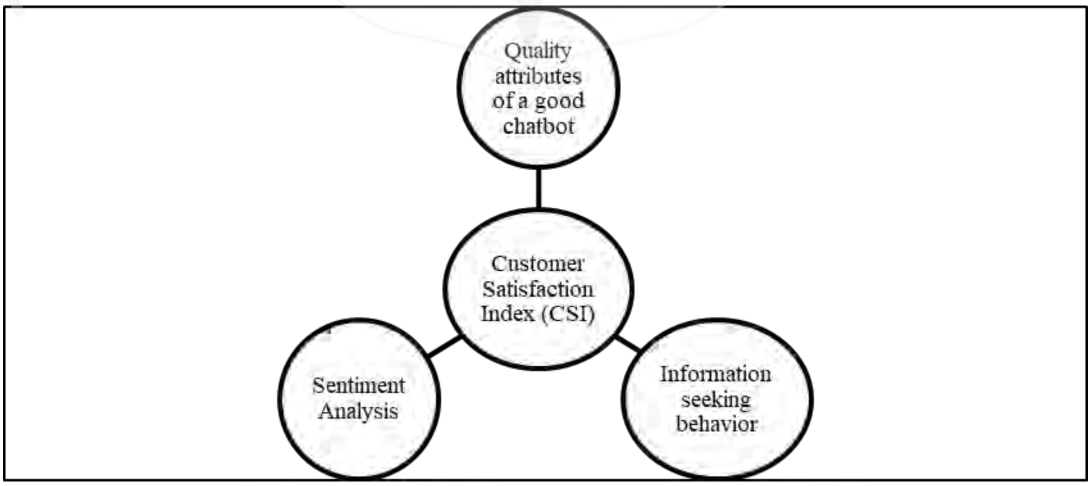


Figure 1.4 Research scope

Table 1.1 Summary for research focus

Research Objectives (RO1 – RO4)	Research Questions (RQ1 – RQ4)	Research Hypothesis (H1 – H2)
• To determine the attributes for a good chatbot • To develop the text analyser using stemming and lemmatization techniques	• What are the attributes for a good quality chatbot? • What are the approaches used to design and develop the new chatbot?	There is no significant difference between features for a good chatbot and customer satisfaction level
• To explore and develop the sentiment from text analysis in chatbot • To validate the user's level of satisfaction with the chatbot's text analyser	• What is the sentiment shown by users from the text analysis of the chatbot? • What is the level of satisfaction shown by user towards the chatbot?	There is no correlation between positive sentiment analysis and customer satisfaction level

1.11 SIGNIFICANCE OF RESEARCH

The research has contributed significantly. The significance of the research can be categorised into four (4) aspects: the quality attributes of a good chatbot; the use of lemmatisation and stemming in the construction of the novel chatbot; the sentiment analysis of the banking users or customers and the level of user or customer satisfaction.

a. Quality attributes of a good chatbot

The identification and determination of the quality attributes is incredibly significant to the research. Numerous studies have been conducted in trying to identify what are the attributes that should be integrated into the design and development of a good chatbot. Many have built chatbots and failed to be effective and satisfactory to the customers. Thus, this study which focused on determining quality attributes is incredibly significant to the research and contributes to future good and effective chatbots.

b. Use of lemmatisation and stemming in chatbot construction

The use of lemmatisation and stemming as pre-processing techniques in the construction of the novel chatbot is incredibly significant to the research. These two techniques helped to reduce words to the base or root form. This means that it reduces the number of unique words in the text conversation made, making it easier for the chatbot to analyse and understand. The stemming would merely remove common suffices from the end of the word tokens, whilst lemmatisation ensures that the output word is in an existing normalised form of a word or lemme (which can be found in the standard dictionary). Thus, using these two techniques is significant to the research. These techniques enhance chatbot's performance and accuracy using Natural Language Processing (NLP).

c. Sentiment analysis in the banking domain

The other significant aspect of the research is the sentiment analysis conducted specifically in the banking domain. Sentiment analysis or what would be normally termed opinion mining, is a Natural Language Processing (NLP) technique used to determine whether data is positive, negative, or neutral. This sentiment analysis conducted in the banking domain, understood significantly; to see how textual data can help the banks monitor brand and product sentiment based on customer feedback and understand customer needs through a transparent and honest chatbot.

d. Customer satisfaction and Customer Satisfaction Index (CSI)

Finally, the other significant aspect of the research is the determination of the level of customer satisfaction or Customer Satisfaction Index (CSI) or the measurement metrics towards the use of the services provided by the chatbot and the organization. The measurement is significant because it is holistic and integrative and does not just measure overall satisfaction. Understanding the customer satisfaction allow any organization to measure the chatbot's performance accurately and allows the organisations to improvise the service to promote better customer engagement.

1.12 CONTRIBUTIONS OF RESEARCH

The research has been able to contribute both theoretically and practically to the field of data mining, data analysis and human-machine interaction. Future researchers can use the findings of the research to enhance the research in both theoretical as well as practical aspects of the field of data mining, data analysis, and human-machine interaction. The contributions can be categorised into the following areas:

a. Theoretical contribution

Future studies can be conducted to compare the techniques used in data mining, Machine Learning (ML) and Natural Language Processing (NLP). This provides a guideline of comparison on various data mining and NLP techniques, highlighting capabilities of different approaches to provide analytical accuracy and effectiveness. This will also help to expand the theoretical understanding and make connection between data mining and NLP with human-machine interaction in specific domains such as banking, healthcare and medicine and education.

b. Practical contribution

Using a practical approach to ensure that the analysis is conducted and implemented at the actual premise related to the domain studied; and technical aspects on energy and visual efficiency elements that have begun can be studied further. All these have contributed tremendously to the research and future researchers can benefit from this. The findings of the research are also useful to future software developers of chatbots that will have a good head start to build an intelligent chatbot that can decipher text and visuals to make the conversation more meaningful and effective and take into consideration elements of energy efficiency and ESG and sustainability at the same time.

c. Applications for organizations

The findings of the research can be a useful contribution to organizations that want their services to be transparent, and honest with good governance and integrity and implement ESG elements not just for the employees of the organization but also their customers.

d. Foundation for future research

This research serves as one of the resources available for future researchers who are interested to widen the matters related to good chatbots and refining aspects on energy and visual efficiency elements in data-driven intelligent systems. Forthcoming studies can also be focusing on the practical execution on these techniques upon various domains such as banking, healthcare, and education.

1.13 LIMITATION OF RESEARCH

The research was conducted with limitations related to aspects such as language limitation, features limited to certain elements, evaluation was conducted in a simulated environment (institution not a real bank), and the samples were randomly selected regardless of the demographic background.

i. The first limitation of the research was based on the language limitation of the chatbot developed. It was designed and developed in the English language which meant that only English language-speaking individuals could be the samples to assess the new chatbot.
ii. The other limitation of the research is the quality features of the chatbot that have been identified and verified. The quality features verified to be included in the chatbot were categorised as technical adequacy, functional adequacy, efficiency, humanity, effectiveness, and ethics. Other than that, the categories were not included in the design and development of the new chatbot. Thus, it is suggested for future researcher to use three (3) or five (5) Round Delphi technique to get input from experts of the field to explore other quality features.
iii. The testing environment was not done at the actual bank environment as it was not allowed for security reasons. Thus, it was limited to a simulation at a university environment in which random samples were bank account holders, regardless of their demographic background.

1.14 DEFINITION OF TERMS

In this research, there are conceptual terms that were stated and discussed. The terms listed are significant for the research design and development of the novel chatbot.

a. Human-Computer interaction (HCI)

Human-Computer interaction or known as HCI is defined as an area that usually used to ease interaction between technology and user (Yun et. al. 2021). Several keys of HCI research are highlighted as usability, accessibility, interaction design, cognitive load, and user-centered design (Dix et al. 2004). For this research the word human-machine interaction was used to mean the same thing as Human-Computer Interaction (HCI).

b. Artificial Intelligence (AI)

Aggarwal et al. (2022) described Artificial Intelligence (AI) as the "study and design of intelligent systems that considered environment and take measure to elevate the chances of success". While McCarthy (2007) defined the term as the science and engineering of making the machine intelligent. For this research, the two definitions were considered but only at the preliminary stage.

c. Sentiment analysis

Sentiment analysis is known as Opinion analysis or Opinion mining (Wankhade et al. 2022). It is a subfield of Natural Language Processing (NLP) which focuses on identifying and extracting information related to emotion (positive, neutral, or negative) from text in different sources of communication platforms such as social media and customer survey. This research had used this definition to conduct sentiment analysis on the use of chatbot with the customers.

d. Energy efficiency

Energy efficiency is defined by many researchers as the saving of energy and use of less energy to conduct a function in an efficient manner. For this research, the development of the chatbot took into consideration on the design of the chatbot in terms of its functionality and the use of certain technology to ensure the user use less power to run the chatbot in an efficient and energy saving way. For example, programming the chatbot to sleep when it is less expected to be used.

e. Customer Satisfaction Index (CSI)

Customer Satisfaction Index (CSI) is a metric used to measure the satisfaction level of users or customers towards services provided by the product or system. In this research, CSI is measured to investigate the perception and satisfaction of users or customers using the newly proposed chatbot.

f. Information Seeking Behaviour (ISB)

Information Seeking Behaviour (ISB) for this research refers to the processes and activities carried out by the users or customers while seeking for information using the newly proposed chatbot. It consists of various strategies, methods and tools to collect and use information from multiple sources such as websites, databases, and expert advice. ISB is represented by several stages which acknowledge the information needed, searching for information through various channels, evaluate the information gathered, and use the information to solve problems or making decision.

1.15 THESIS ORGANISATION

This thesis was organised into six (6) chapters as indicated below:

Table 1.2 Thesis organisation

Chapter	Description
Chapter I	**Introduction** This chapter provides the research background and describes the research focus. Other elements such as the research objectives, research questions and research hypothesis were also stated clearly. The last section of this chapter discussed the significance of the research and its limitations.
Chapter II	**Literature Review** This chapter is a particularly important chapter for the thesis and comprises a review of past research from both primary and secondary sources related to the research problem. The chapter will guide the flow and the expected process that will be discussed in the following chapters.
Chapter III	**Research Methodology** This chapter elaborates on the method used to conduct the research based on the research framework and the approaches used to support and ensure the success of the research conducted.

to be continued…

…continuation

Chapter IV	**Design and Development**
	This chapter discusses on the processes involved in designing and developing the new chatbot which includes the new dimensions such as the framework, techniques, and approaches, as well as the evaluation and testing of the new chatbot.
Chapter V	**Results and Discussion**
	This chapter discusses on the outcomes of the data to answer the research objectives (RO1- RO4), research questions (RQ1-RQ4), and research hypotheses (H1-H2). Data is displayed in both tables and graphs to ease understanding. Discussions on finding of the sentiment analysis and customer satisfaction are also discussed in this chapter.
Chapter VI	**Recommendations and Conclusion**
	This concluding chapter discusses on the conclusion and recommendations for future research (both theoretical and practical aspects of the field) that can be conducted by the future researchers in the field.

1.16 CONCLUSION

The evolution of the information-searching platform allows vast technological enhancement to assist the task assigned. The history of chatbots goes back to the 1960s when Artificial Intelligence (AI) was created to assist humans with their daily tasks. This chapter laid out the research focus which is research objectives, research questions, and scope. The details for each area are mapped in the research theoretical framework and conceptual framework with support from related theories and literature. The chapter wrapped up with the significance of the research and the limitations observed from this research. The concluding section listed the thesis organization which explains items discussed in each chapter.

CHAPTER II

LITERATURE REVIEW

2.1 INTRODUCTION

This chapter discusses and emphasizes the critical review of previous research conducted which are related to the current research. The history and state-of-the-art of the chatbots, aspects on sentiment analysis in human interaction, as well as the customer satisfaction index (CSI) will also be reviewed and discussed in this chapter.

2.2 HISTORY OF CHATBOTS

Chatbots are conversational tools that have become a rising trend, using instant messaging as a platform. It is a machine agent that acts as a natural language user interface to provide data and services (Dale 2016) to help people get through routine tasks quickly and efficiently. It is used by developers for mobile messaging applications (Folstad & Brandtzaeg 2017) designed to simulate human-like conversations via SMS. Oracle (2016) reported that 80% of businesses use or plan to use chatbots. It is estimated that the market was worth $USD2.6 billion in 2020 and is expected to reach more than $USD10.0 billion by 2025.

The history of the chatbot started in the 1950s with ‘The Turing Test’ by Alan Turing. The test aimed to decide whether a machine has similar human-like intelligence or the vice versa (Vogel 2017). Initially, it was conducted as a game where humans interacted with two parties through textual messages. The respondents needed to guess which of the two parties were machines. If they could not guess correctly, the machine passed the Turing Test.

Then, in the 1960s, ELIZA was released. ELIZA is a bot using an early Natural Language Processing (NLP) program, created by Joseph Weizenbaum. It uses a simple pattern cross-checking and a template-based response mechanism to match the conversational style of a non-directional psychotherapist (Weizenbaum 1966). This system imitates human conversations by cross-linked user input with the scripted knowledge base. Rogers (1995) through his works, found that the non-psychotherapist known as Rogerian psychotherapy is a 'person-centred therapy' that aims to inform clients of their attitudes and the way they behave. In 1988, Jabberwacky introduced a bot using this 'person-centred therapy' approach which became the earliest creation that attempts to promote interaction between humans and machines. In 1992, Dr Sbaitso created an artificial intelligence program or system for MS-DOS-based personal computers distributed with various sound cards and acts to demonstrate a digitised voice (Pavlíček 2021).

Artificial Linguistic Internet Computer Entity (A.L.I.C.E) is a Natural Language Processing (NLP) bot. ALICE referred to as Alicebot uses heuristic pattern matching rules to human input that resonance with the conversation. Thompson (2002) stated that A.L.I.C.E, which was released in 2001 has gained popularity for its almost realistic behaviour. In 2010, IBM Watson was released. It uses Natural Language Processing (NLP) and Machine Learning (ML). Watson sparked curiosity around a 'machine that could think' and opened possibilities that could be done with Artificial Intelligence (AI). Watson currently, is applied in most business services (medical/health, finance, and the like) to unlock insights, drive productivity, and deliver better customer experience. Apple iOS has integrated Artificial Intelligence (AI) elements in their devices which is known as Siri in 2011. It is intelligent yet bold and uses a natural language user interface to answer inquiries and perform web service requests. Vogel (2017) in his work, explained that users are more aware of Siri as the functionality has improved over the years. The popularity of Siri grows with the fact that it learns what the users' needs are.

Google has also stepped up their strategy in the Artificial Intelligence (AI) field when it introduced Google Now in 2012. It was developed for the Google Search mobile app. It uses NLP to answer questions, make suggestions, as well as lead users to a set of web services. In 2015, Alexa was Amazon's voice AI service planted into the

Amazon Echo device. It can interact with humans using an NLP algorithm to receive, recognise, and respond to voice instructions. In the same year, Cortana was created as an intelligent personal assistant by Microsoft. Fundamental Cortana can help users access information quicker, connect with other people faster, and be good at keeping on track. Thus, it serves to set reminders, recognise voice instructions, and use Bing as a search engine to answer questions from users. In April 2016, Facebook launched an intelligent bot using the Messenger platform. It can interact with Facebook users through keyword recognition and pre-programmed replies. Currently, there are more than 11,000 bots available.

Figure 2.1 explains the general chatbot architecture which consists of five (5) main components: user interface, user message analysis, dialogue management, backend, and response generation (Adamopoulou & Moussiades 2020).

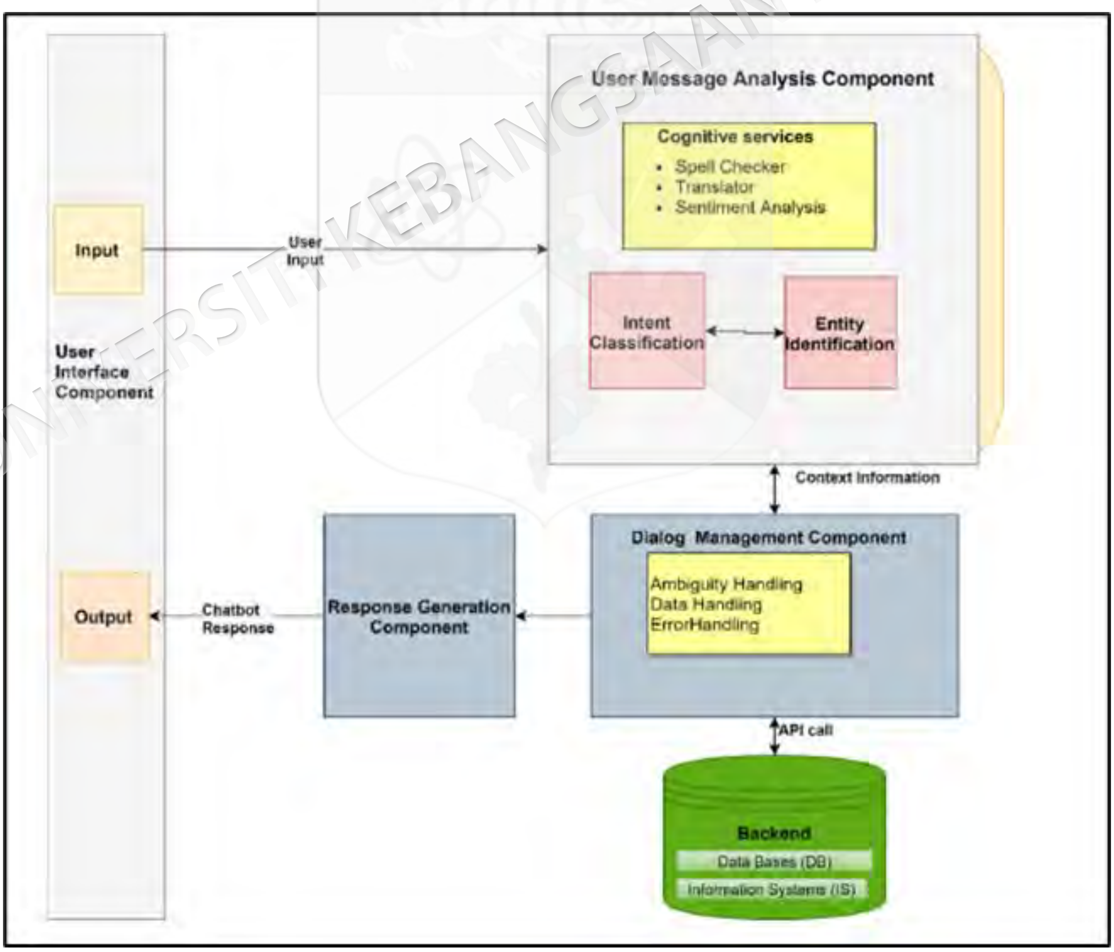


Figure 2.1 General chatbot architecture

Source: Adapted from Adamopoulou & Moussiades (2020)

The user interface component allows users or customers to communicate and interact with the chatbot through messenger applications such as Facebook Messenger, Eliza, Cortana, and the like. The operation of the chatbot would begin with the users' or customers' requests and queries. The user message analysis component would extract the information based on the users' or customers' requests or queries and produce a representation of its meaning that could be used later in the process based on natural language understanding. The dialogue management component deals with information that comes from other components and is responsible for controlling and updating the context of conversations between the users or customers with the chatbot and at the same time, governing the actions of the chatbot. The next component, which is the backend component, would retrieve information that is needed for performing the required tasks, and then forward the message to the dialogue management component and the response generation component. The information is stored digitally in large repositories. Finally, the response generation component will ensure that the appropriate information that has been retrieved will determine the content of the best approach to respond in express to the users or customers. Thus, the response generation component responds in a more understandable format for the users or customers.

2.2.1 Implications of History of Chatbots to Current Study

The evolution of chatbots from year to year parallels the need of users of customers to solve their daily needs. Starting from assessing the interaction between humans and machines escalated to the development of chatbots to become intelligent assistance to humans with ample training and knowledge that fed into the knowledge base. From the chatbot architecture design by Adamopoulou and Moussiades (2020), the elements presented must be considered before designing and building the desired chatbot.

AI-based system such as Siri, Google Assistant and Amazon's Alexa has grown initially from the ELIZA which nowadays is seen as proof for successful integration of Natural Language Processing (NLP) and Machine Learning (ML). This integration allows to improvise user's experience by facilitating time taken for them to resolve their difficulties.

The usage of chatbots in banking environment have progressed from simple automated feedback system to progressive Artificial Intelligence (AI) based virtual assistant that are skilled to provide personalised assistance and advice, processing transactions, and all-time (24/7) available customer service.

The history of chatbots certainly has implications to the general architecture of chatbots. The general architecture of chatbot that consists of user interface, message analysis, dialogue management, backend, and response generation has become the backbone for chatbot development of the current study, with aim to manage user tasks and provide precise feedback. This architecture that has expanded to cater this study is needs are reflected in the design and development of the newly proposed chatbot.

Governing the history into the foundation of designing and developing chatbot for this research meets the objective of this research to contribute to the field of Artificial Intelligence (AI) and customer experience especially in the Banking domain. The proposed chatbot for this research designed in form of interactive ways to promote beneficial conversation and meet user's or customer's expectation employs the humanity approach. By adapting the architecture from previous designed chatbot as mentioned in past literatures, the new proposed chatbot as mentioned in Chapter III will then to answer the research questions mentioned in Chapter I.

2.3 CLASSIFICATION OF CHATBOT

There are four (4) identified main categories of chatbot applications namely advisory, entertainment, commercial as well as service chatbots (Bayu & Ferry 2016). The advisory chatbots provide advisory services or recommendations but do not act to enforce the suggested matters to the users or customers. This service aims to offer maintenance and provide recommendations on services as well as suggestions (Soufyane et al. 2021). The entertainment chatbots serve to engage users or customers by providing information related to entertainment programs or shows, on upcoming events, as well as assisting in ticket bookings. Commercial chatbots focus on easing the user's or customers' experiences in purchasing tickets and other items. Lastly, the service chatbots focus on providing facilities to the desired users or customers (Soufyane et al. 2021).

These four (4) types of chatbot applications have also been known as response generation-based, service-based, knowledge-based, and goal-based. Table 2.1 summarises the mentioned chatbot categories.

Table 2.1 Classification of chatbot application

Types of Chatbot Application	Details
Goal-based	• Activity-based • Conversational • Informative
Response generation-based	• Retrieval-based model • Generative model
Knowledge-based	• Open domain • Closed domain
Service-based	• Inter-personal • Intra-personal • Inter-agent

Source: Soufyane et al. 2021

Borah et al. (2019) listed the several types of models for chatbot design and development. It consists of retrieval-based, generative-based, long and short conversations, as well as open and closed domains. Retrieval-based models utilise the resources from predefined responses and choose suitable responses according to the input and context of the conversations. This model guided the end user to follow the interaction flow of the script provided as most of the chatbots are available in a structured and scripted form. This model deals with one domain with less flexibility. The generative-based model is based on Machine Translation techniques which translates from an input to an output response. It is also known as Artificial Intelligence (AI) Chatbot. The key features listed are the successful rate of the implementation and the ability to manage more than one domain. This model is self-trained using the substantial amounts of data received within the interaction. However, the downside of this model is the cost and the complexity of the model. The third type is the long and short conversation model. This model can divide the conversations into long and short portions. It is known that the longer the conversations, the more complex the automation is, as it is crucial to keep track of the inputs provided throughout the conversations which is contrary to the short conversations. The last model is the open and closed domain model. The open domain model does not have a constraint to any goal or intention, which requires an infinite range of knowledge to provide acceptable responses. However, the close domain model assigns limitations to the possible input and response as the system is designed for a particular objective.

2.3.1 Implications of Classification of Chatbots to Current Study

Past research on the classification of chatbots function as the benchmark towards design and development of current chatbots. The classification stated in the Table 2.1 was based on application oriented which conduct different tasks to alleviate customer's experience while using service provided by the chatbots. The functionalities application of chatbots are widespread across the field such as advisory, entertainment, commercial and service chatbot. These categories are further advanced into sub-categories based on type of responses produce which are goal-based, response generation-based, knowledge-based, and service-based models.

The classification of chatbots has significant implications to the current study. The current study designed and developed a chatbot serves as an advisory chatbot. It aims to provide information related to services provided by the banking institution. The assistance provided includes provide recommendations on banking products and explanation on current products provided by the institution. The effectiveness depends on the quality and relevance of the information provided throughout the conversations with focus on the ability to understand the user context and preference. With respect to past literature recorded by Soufyane et al. (2021), the proposed chatbot in this study can be referred as goal-based chatbot. The goals designed for the new chatbot developed in this study aims on task-oriented which during testing, the targeted users were given a set of tasks to be evaluated with goal stated. Amendments were made based on the feedback provided by the users lead to development of improvised chatbot version which was evaluated in the second phase.

2.4 CONVERSATIONAL ASPECTS OF CHATBOTS

Chatbots are built with a concept of conversations between two people in the mind of the user as well as the developer. Any human-system conversations consist of two (2) parties: the user (the one who looking for information) and the system itself which acts as the expert (Duijst 2017). Quarteroni and Manandhar (2009), classified aspects and issues that are related to human conversations. The first aspect is context-maintenance which refers to making use of conversations to interpret the user's input. The second aspect is expressing understanding to follow up and clarify the previous conversations. Next is a follow-up proposal that aims to motivate users or customers to give feedback on the answers given whether they are satisfactory or vice versa. The last aspect is the natural interaction which would create a smooth and active dialogue sessions between the chatbots and the users or customers.

2.4.1 Implications of Conversational Aspects of Chatbots to Current Study

Creating a chatbot with natural interaction features allow users to feel at ease while communicating to get their tasks resolved. This is reflected through the satisfaction level shown after using the services which are part of the items of interest to be determined at the end of the chatbot testing in this study for the new chatbot developed.

The newly proposed chatbot applied conversational aspects as parts of its persona to portray the closest interaction as human interaction with users of customers. In this study, the new chatbot developed emphasises on improving the conversational facility, context-wise, sentiment analysis, and user satisfaction. The newly proposed chatbot incorporates Natural Language Processing (NLP) techniques, Machine Learning (ML) algorithm and sentiment analysis to better understand user's query and expectation in using banking services.

Another impact of conversational aspect in the newly proposed chatbot, it helps researcher to determine the users or customer intention while using the system. During the preliminary and post-development testing, most customers were interested to know the main services provided by the banking system such as apply for loan, credit card as well as housing loan. At the end of each testing session, they were asked to provide any suggestions or comments for improvement which later helps to change to better acceptance during the final testing. Serban et. al. (2017), stated that involvement of Natural Language Understanding (NLU), to recognize the intention allows chatbot to identify various users' needs and respond respectively.

2.5 HUMAN-COMPUTER INTERACTION (HCI)

Human-Computer Interaction (HCI) is defined as the multidisciplinary study focusing on developing designs, conducting user requirement studies, evaluate interactive computing systems, and aims to improve usability and user experience (MacKenzie 2024). The information passes between user, and the system is known as the loop of interaction (Aronson et.al. 2005; Newell & Simon 1972). The information consists of task environment, machine environment, areas of interface, input flow, output, and feedback. Task environments focus on the rules and goals determined by the user. Machine environments provide space that connect to the machine. Areas of interface is a process that focus on human and computer (non-overlapping areas) and between users (overlapping areas). Input flow focus on the flow of information started in the task environment.

There are several key aspects of human-computer interaction in chatbot such as: Natural Language Processing (NLP), User Centered Design (UCD), conversation flow and context management, feedback and affordances, personality and emotional intelligence and multimodal interaction. NLP in chatbot allows the system to understand and produce human-like interaction with users or customer. While UCD focusing on designing chatbot based on the requirements, behaviours and constraint of the end users or customers to make certain on usability and accessibility (Norman 2013). An effective chatbot able to maintain the conversation context, support consistent and uninterrupted conversation. Chatbots need to provide end users or customers with unambiguous feedback as indicator and confirmation to guide the conversation flow. Personality of chatbot also plays an important role in designing a chatbot. This allows the chatbot to be more interactive mimicking human engagement which then improve user's experiences (Bradeško & Mladenić 2012). Lastly, the multimodal interaction which comprises of voice, text, and graphic to elevate accessibility and effectiveness of chatbot also depicts the human-computer interaction attributes.

2.5.1 Implications of Human-Computer Interaction to Current Study

Human-Computer Interaction (HCI) plays a big role in this study at the designing, development, and testing phase. The objectives are to cater system usability and support overall user experience. Over the years, HCI continue to evolve parallel with the integration of the usability principles and user-centered design. With advancement of Artificial Intelligence (AI) and Machine Learning (ML), this allows and support natural, flexible, and intuitive interaction with the newly proposed chatbot.

The newly proposed chatbot's interface was developed with elements such as simplicity, direct, user-centered, and user-friendly. The use of avatar and interactive colours can retain user attention while using the service. Interaction flows which mimics human real interaction provide ease of communication throughout the session allows NURUL-C to achieve the aims of its development. User's feedback was recorded at the end of the session in the form of satisfaction level ratings which help in consideration for amendments.

2.6 QUALITY ATTRIBUTES CHATBOTS

High-quality software is the most important aspect achieved by any software developer (Pressman 2010). A standard design convention widely used in any software development process is known as ISO 9126. It is a standard guideline for developers that defines quality as the sum of features or characteristics that depend on the ability to satisfy stated and indirect needs. ISO 9126-1 defined (6) characteristics for an overall quality software model which are: functionality, reliability, usability, efficiency, maintainability, and portability followed by other twenty-two (22) sub-characteristics as explained in Table 2.2. This model is an international standard that is accepted and used for software evaluation globally.

Table 2.2 Characteristics of the ISO 9126-1 Quality Model

Characteristics	**Description**	**Sub-Characteristics**
Functionality	The capability of any software to provide service or solution to users to meet the intended needs and conditions.	• Suitability • Accuracy • Interoperability • Compliance • Security
Reliability	The ability of the software to work in any environment or condition.	• Maturity • Fault tolerance • Recoverability

to be continued…

…continuation

Usability	Ease of use which covers user-friendly, cosmetic adequacy, easy to learn and simple navigation approach.	• Understandability • Learnability • Operability • Attractiveness
Efficiency	Represented by time taken to complete any given task to the system.	• Time Behaviour • Resource Utilisation
Maintainability	Easy to maintain in terms of adding new commands or rules to the existing system; able to upgrade to latest version at any time; and maintenance and testing should be easy and cost-effective.	• Analysability • Changeability • Stability • Testability
Portability	The ability to move across the environment.	• Adaptability • Installability • Conformance • Replaceability

In research conducted by Smutny and Schreiberova (2020) on forty-seven (47) chatbots that were suitable for learning, it recorded that there were four (4) main categories for chatbot quality assessment. It listed teaching, humanity, affect, and accessibility. Table 2.3 describes the details for each of the categories.

In another research conducted by Meerschman and Verkeyn (2019), the researchers have listed and categorised all features from past literature into twenty-eight (28) attributes. The selection of literature is based on the objective of obtaining a list that will enable users and the chatbot builder to evaluate attributes of chatbots as thoroughly as possible to find the best quality attributes. Table 2.4 shows the list of quality attributes across research found between 1966 and 2019.

Table 2.3 Quality attributes for chatbot assessment

Category	**Quality Attributes**	**Examples**
Teaching	• Recommend learning content. • Provide feedback, Q&A • Set goals and monitoring learning progress.	• Links to web pages or documents with learning topics. • Quizzes and tests provide instant feedback on each question once answered. • Quizzes and tests are counting scores. • Ability to resume test / quiz and continue afterwards.

to be continued…

…continuation

Humanity	• Able to maintain the theme of the discussion. • Able to respond to specific inquiry.	• Chatbot able to interpret command given correctly. • Portray natural and convincing interactions. • Flexible interpretation of knowledge.
Affect	• Provide greetings and have pleasing personality. • Entertaining, engaging.	• Offers greetings such as “Hello” with the user’s name. • Chatbot used jokes, humours, emoji, animated GIFs from current trend of conversations.
Accessibility	• Able to pick up the meaning and intent from the conversation. • Respond to social cues accordingly.	• Able to provide respond to the question “Sorry, I don’t speak English” or “Sorry, could you repeat the question?” • Read and gives feedback towards users’ or customers’ emotions.

Source: Smutny & Schreiberova 2020

Table 2.4 Chatbot quality attributes

Dimension	**No**	**Quality Attributes**	**Reference**
Functionality	1	Interpret commands accurately	Morrisey and Kirakowski (2013)
	2	Execute requested tasks	Explorative study
	3	Flexible in interpreting knowledge	Cohen and Lane (2016)
	4	Able to maintain a discussion	Morrisey and Kirakowski (2013), Kuligowska (2015)
	5	Activation	Hertzum et.al. (2002), Kuligowska (2015)
	6	Number of services available in the chatbot	Eeuwen (2017)
Trustworthiness	7	Containing dependable / accurate information	Explorative study
	8	Possibility of rating the Chatbot	Kuligowska (2015)
	9	Contains wide range of knowledge	Cohen and Lane (2016), Kuligowska (2015)
	10	Robustness to unexpected input	Kuligowska (2015)
	11	Transparency	Hertzum et. al. (2002)
Safety/Intrusion	12	Protect and respect privacy	Eeuwen (2017)
	13	Save from intrusion	Duijst 92018)
Efficiency	14	Ease of use	Candela (2018), Duijist (2018)
	15	Quick replies to vs free text	Duijst (2018)
	16	Available always	Wang et. al. (2019)
	17	Accessibility	Duijst (2018)
	18	Need of an account	Explorative study

to be continued…

…continuation

Graphical Appearance	19	User-interface	Duijist (2018), Kuligowska (2015)
	20	Use of emojis and pictures / gifs	Kuligowska (2015)
Humanity	21	Realness of the chatbot	Ramos (2017), Wallace (2003), Weizenbaum (1966), Cohen (2005), Whitby (1996)
	22	Create an enjoyable interaction	Morrisey and Kirakowski (2013)
	23	Convey positive personality	Morrisey and Kirakowski (2013), Kuligowska (2015)
	24	Read and respond to moods	Beale and Creed (2009)
Empathy	25	Personalisation options	Kuligowska (2015)
	26	Personalised suggestions	Explorative study
Responsiveness	27	Responding immediately	Explorative study
	28	Productivity	Rieke (2018), Candela (2018)

In another separate research conducted by Meerschman and Verkeyn (2019), they discovered a concept called the Analytical Hierarchy Process, introduced as a guidance for the decomposition of unstructured situations into a simple version for decision-making (Smutny & Scheriberova 2020). These were divided into seven (7) main dimensions and verified through the Three-Round Delphi Technique by experts in the field. These dimensions are namely: technical functionalities, efficiency, effectiveness, humanistic, ethics, technical satisfaction, and accessibility. Table 2.5 summarises each dimension which was adapted from Saaty (1990) who first introduced the Analytical Hierarchy Process.

Table 2.5 Plausible quality attributes of the chatbots

Dimensions	**Attributes**
Technical Functionalities	• Sentiment analytics: Intelligent (Using AI, NLP, ML, Semantic technology) • User friendly • Automated • Simple interface • Interactive (using video, audio) to cater needs from different group. • Multi-lingual • General ease of use
Efficiency	• Robust to manipulation of data input by users • Provide proper escalation channels • Quick answer • Can perform damage control easily

to be continued…

…continuation

Effectiveness	• Interpret statements and instructions accurately • Linguistic accuracy • Execute desired task correctly • Consists wide array of knowledge • Able to perform simple problem solving
Humanistic	• Human-like personality • Transparent and disclose the chatbot identity. • Able to respond correctly to user's or customer's queries • Can maintain the pattern of conversation
Ethics	• Trained with knowledge cultural and ethics of users • Protect and respect user's or customer's privacy • Sensitivity to user's or customer's concern (i.e security, social concerns) • Constancy
Technical Satisfaction	• Able to convey greetings • Entertaining, fun, and users enjoy the conversation • Able to respond to user's mood • Provide emotional information using tones and expression
Accessibility	• Ability to detect user's intent • Meeting with vast range of needs (i.e time of buffer, text interfaces) • Available 24/7

Source: Adapted from Saaty (1990)

2.6.1 Implications of Quality Attributes to Current Study

Researchers found focusing on the attributes that a chatbot should have to promote easiness of use and satisfied experience. Among key attributes mentioned in the literature gathered was usability. This element is crucial to ensure end users or customers can interact with the chatbot without any setbacks or frustration. Spontaneous flow, simple language, and efficient error handling can promote to better user's engagement. Følstad and Brandtzæg (2017) mentioned that "a usable chatbot must be able to support clear conversation flows, provide real-time response, and handle any misunderstanding gracefully".

From the literature collected, the prominent attributes selected for the development of NURUL-C was functionality, efficiency, technical satisfaction, humanity, effectiveness, and ethics which added in the later phase with attributes on visual element and energy efficiency. These listed attributes guide the chatbot development and amendment process to gain data on sentiment used and use by end users or customers.

2.7 BENEFITS USING CHATBOT

Chatbots have made vast contributions in various fields. In a study by Hill and Farreras (2015), they found that interaction between chatbots and humans tends to last longer than interaction among humans particularly between strangers. Human-chatbot interaction included shorter messages and less complex vocabulary. James (2016) believes chatbots can be accepted as helpers that aim to assist users in exploring the content or services available online.

Aishwarya et al. (2017) have listed the significance of using chatbots in e-commerce. Firstly, the chatbot supports customer engagement by learning from the interaction, allowing it to send related information about various brands, products, and services. Secondly, a chatbot can collect customer feedback and data through natural conversations conducted with customers. Besides that, the chatbot helps categorise items from e-commerce websites according to respective groups and can control the website's traffic. Xu et al. (2017), suggest that chatbots are a promising substitute for customer service whereby customer finds that chatbots can help them obtain answers to questions, provide suggestions for purchases, make orders, and provide updates on shipping status using a natural language interface.

2.7.1 Implications on Benefits of Chatbots to Current Study

Using chatbot provides abundant benefits towards the organisation that incorporate these systems into their services as well as towards users or customers. Availability is the key to any service which chatbots provide round-the-clock service which users or customers can get their information of making decision at any time. Chatbots can provide instant feedback towards Frequently Asked Questions (FAQ) which helps reduce in waiting times. The usage of this system also helps in cost reduction on human workforce in which chatbots help to manage repetitive task handle by more than one agent. This system also promotes customer satisfaction when users or customers need to spend less time to get response and consistency of response providing help to avoid misinterpretation.

The newly proposed chatbot is expected to contribute to banking domain to help customers support team to provided initial information to users or customers who are looking for instant response especially during 'out-of-office' hours. Real-time data collection from users or customers helps the organization to collect customer preference, pain points, and suggestions for improvement. The newly proposed chatbot is also expected to promote users' or customers' engagement through attractive interface and interaction experience which will make them feel connected.

2.8 WEAKNESSES AND THREATS OF CHATBOTS

Although there are many positive contributions of chatbots towards human daily lives, there are still gaps in the performance of chatbots. Verstegen (2022) shared various weaknesses of chatbots. For example, the chatbot cannot understand human natural language, which includes slang, incorrect spelling, and double-meaning words. Whenever the chatbot cannot detect these categories of words, it will impact the response rates, which then may lead to frustration for the users or customers. Another weakness is that the chatbot has limited response capability, which might leave users or customers with incorrect or no solution at all. Chatbots are said to be unable to answer multiple questions that require them to reason and thus assist in decision-making unless they are highly intelligent (this is of course not impossible if advanced AI is applied).

Other challenges addressed is the capability to handle complex problems. It is proven that chatbots can solve simple and repetitive tasks. However, the drawbacks are set on multiple emotional express through users or customers' requests. The impact seen was on the users' or customers frustration on the mismanagement caused (Van der Heijden 2020).

Trust and technology adoption by users or customers is also part of the challenges addressed in implementing chatbots into the workflow. Insecurity surface as users or customers may be reluctant to share their sensitive information especially when the system does not provide any human support intervention for further assistance (Lase 2020).

2.8.1 Implication of Weaknesses and Threat of Chatbot to Current Study

Security of information input is one of the biggest factors that must be considered during the development and system execution stage. As mentioned by Radziwill and Benton (2017), a conversational agent must be designed with strong security measures to prevent personal data leaking. For the newly proposed chatbot, it will deal with less sensitive information, as it only provides services related to banking products. Thus, it will impose less threat on personal data leakage.

2.9 CHATBOT EVALUATION FRAMEWORK

Russell-Rose (2017) found that five (5) types of perspectives were derived from the chatbot attributes: namely user experiences, information retrieval, linguistic, technology, and business. Table 2.6 summarises the plausible perspectives regarding the attributes of chatbots (Peras 2018), that could encompass the chatbot evaluation framework.

Table 2.6 The chatbot evaluation framework

Perspective	Category	Attributes	Metrics	Approach
User Perspective	Usability	• Completion of task • Receiving assistance or information • Support minimal set of commands • Response type frequency	• Response type relative frequency • Percentage of match found • Response type based on relative probability • Rating scale • Surveys • Questionnaires • Assistance from Help and Cancel commands	Qualitative, Quantitative
	Performance	• Robustness • Response in unexpected situations • Consistency • Effective task allocation	• Percentage of successes • Rating scale	Qualitative
	Affect	• Personality • Emotional information • Entertainment • Involvement • Personality traits • Human assistance provision • Reliability • Privacy protection	• Rating scale • Surveys • Questionnaires • Checking for keywords • Number of dialogues turns • Total conversation duration	Qualitative, Quantitative
	Satisfaction	• Expectation • Impression • Command • Navigability • Engagement • Entertainment • Curiosity • Social relations	• Conversation duration • Number of conversations turns • Rating scale	Qualitative, Quantitative

to be continued…

…continuation

		• Ability to learn • Ability to aid		
Information retrieval perspective	Accuracy	• Ability to foresee language variations	• Precision • Recall • Typing errors and synonyms	Quantitative
	Accessibility	• Ability to detect meaning and intent and able to respond appropriately	• Context sensitiveness • Percentage of successes • Number of unfitting responses • Turn correction ratio	Quantitative
	Efficiency	• The use of resources to achieve the goals.	• Matching types • Measuring the answer time for the commands and obtaining mean values • Total elapsed time • Total number of users turns. • Total number of turns per task	Quantitative
Linguistic perspective	Quality	• Accuracy of the responses • Categorization of responses	• Likert scale	Qualitative
	Quantity	• Ampleness of information	• Likert scale	Qualitative
	Relation	• Relevancy of responses to the context of the conversations	• Likert scale	Qualitative
	Manner	• Ambiguity of the responses	• Likert scale	Qualitative
	Grammatical accuracy	• Acceptability from grammatical and meaning perspectives	• Total number of errors made in the chat period. • Word-level analysis (vocabulary range, spelling, upper/lower case) • Grammar-level analysis (nouns, pronouns, verbs, word order, and others)	Qualitative

to be continued…

…continuation

Technology perspective	Humanity	• Naturalness • Maintaining themed discussions • Responding to specific questions • Non-understanding rate	• Turing Test • Rating scale • Percentage of success • Percentage of rejection	Qualitative, Quantitative
Business perspective	Business value	• Efficiency cost • Qualitative cost	• Number of users • Duration of the chatbot conversation • Number of conversations made by chatbot. • Number of agents involved in the conversation. • Duration of the conversation with agent • Number of failed conversations • Number of inappropriate responses • Number of queries repeated	

Source: Adapted from Peras (2018)

Goli et al. (2023), conducted a study to investigate the impact of factors, namely the perceived ease of use, perceived usefulness, perceived enjoyment, innovativeness, perceived information quality, as well as perceived customisation towards the behavioural intention to use chatbots. The researchers designed an empirically validated model using structural equation modelling with support from the AMOS software. The study used a five-Likert Scale for 378 chatbot users using an online approach. The outcome from the survey showed that the six (6) attributes selected, which were perceived ease of use, perceived usefulness, innovativeness, perceived information quality, and perceived customisation, provided positive impacts toward the intention of using chatbots compared to perceived enjoyment, which showed no significant effect. Table 2.7 shows the attributes and statements compiled and used throughout the study.

Table 2.7 Operationalisation of construct

Construct	**Items**	**Reference**
Perceived ease of use	• My interaction with the chatbot is clear and understandable. • My interaction with the chatbot does not require any mental effort. • It is easy to become skilful at using chatbot. • Overall, I find chatbot is easy to use.	Davis (1989)
Perceived usefulness	• Chatbot app provides useful information and service to me. • Chatbot app helps me to find information quickly. • I think chatbot enhance the effectiveness of my life in general. • Overall, I find chatbot app is useful to me.	Davis (1989)
Perceived enjoyment	• It is fun using chatbot. • Chatbot provides me with a lot of enjoyment. • Using chatbot app is unpleasant	Venkatesh et al. (2002)
Innovativeness	• If I knew about chatbot, I would look for ways to experiment with it. • Among my friends, I am the first one to experiment with the use of chatbot. • I like to experiment with chatbot. • I would try chatbot even my friends have not tried it before	Karaiskos et al. (2009); Aldas-Manzano et al. (2009)
Perceived information quality	• The information provided by chatbots is accurate.	Doll et al. (1994); Nelson et al. (2005)

to be continued…

…continuation

	• The information provided by chatbots is complete. • The information provided by chatbot is concise. • The information provided by chatbot is useful. • The information provided by chatbot is comparable to other outputs. • The information provided by chatbot is easily understood	
Perceived customization	• The message/ information provided by chatbot is tailored to my needs. • This information/message provided by chatbot makes me feel that I am a unique customer. • I believe that this information/message provided by chatbot is customised to my needs.	Gazley et al. (2015)
Behavioural intention to use	• I expect to use chatbots for networking. • I am likely to use chatbot. • I would continue to use chatbot.	Bhattacherjee (2001)

Source: Adapted from Goli et al. (2023)

2.9.1 Implications of Chatbot Evaluation to Current Study

Evaluation chatbot is needed to determine the efficiency of the system to deliver information and help in decision making. Various impact factors listed in past studies functioned as guideline for building an effective and user-friendly chatbot.

Implication of applying chatbot evaluation into the system development and deliverable can be seen in different areas such as user experience and satisfaction, performance and effectiveness, continuous improvement, as well as ethical and social considerations. User Experience (UX) while using the chatbot can help to identify any usability difficulties and guide the process of modification. The smoothness of interaction with the chatbot indicates the efficiency of the system provided to the users or customers. UX also helps the developer to manage and tackle the error or misunderstanding caused while using the system to avoid outbreak. Other direct impact of chatbot evaluation is promoting better and accurate response towards queries. Continuous improvement is needed to train the system to provide accurate response and meet the users' or customers' expectations in spans of time aligned with the technology and business advancement. Ethical aspect is part of the evaluation focus as it involves exposure of users or customers to their data privacy and security. This can promote trust among users or customers to use the system.

The newly proposed chatbot would be built and evaluate based on the factors such as user experience, easiness of information retrieval and linguistic efficiency. Application of both user satisfaction (CSAT) and sentiment analysis guiding this study to achieve its fourth objective. In CHAPTER V, user's satisfaction level would be measured based on the time taken for the newly proposed chatbot and accuracy of response provided. In the initial phase of testing, most users or customers leans toward neutral response which would indicate for improvement to be made on the system. During the next phase of testing, improvement on the feedback received can be seen with hopefully, positive feedback provided for each test cases provided.

2.10 NATURAL LANGUAGE PROCESSING (NLP)

Natural Language Processing (NLP) is a machine-learning technology that gives computers the ability to interpret, manipulate and comprehend human language. NLP is a subfield of AI and helps machines in this case chatbots to understand human language so that they can automatically perform repetitive tasks such as summarisation, translation, ticket clarification and spell check. Thus, NLP is also a technique known for "text mining, machine interpretation, as well as mechanised inquiry" (Yaakub et al. 2019). It is a technique used to help understand and manipulate natural language in both text and language form to meet the purpose of the task given. Li (2017) described advantages and challenges of deep learning for NLP that can be observed in Table 2.8.

Table 2.8 Advantages and challenges of deep learning for NLP

Advantages	Challenges
• Good at pattern recognition problems	• Not good at inference and making complex decisions
• Data-driven and high performance in most problems	• Unable to oversee symbols
• End-to-end training: with little or no domain knowledge needed in system construction	• Require vast collection of data
• Learn of representations: cross-modal processing is possible	• Difficult to manage long tail phenomena

to be continued…

…continuation

• Gradient-based learning: learning algorithm is simple • Mainly supervised learning methods	• Model is usually a black box and hard to understand • Computational cost of learning is high • Unsupervised learning methods need to be developed • Still lacking in theoretical foundation

Source: Li 2017

Divya et al. (2020) have also found advantages and disadvantages of Natural Language Processing (NLP). Table 2.9 summarises the advantages and disadvantages.

Table 2.9 Advantages and disadvantages of NLP

Advantages	Disadvantages
• NLP incorporates the automatic summary generator which offers a readable summary of text. Newspaper is the example for automatic content generator.	• The actual meaning failed to convey due to context expert knowledge such as sarcasm.
• Assist to determine which words are in separate portion from the main text referring to the same object whenever there are multiple sentences. The process is called Co-reference Resolution (CR).	• Unable to process visual context
• Adopted by various organisation to improve the efficiency of documentation, which is reflected on the accuracy and able to find related information from the large unit of databases.	• Unable to understand generalised searched and difficulties to deal with synonyms.
• Important level of flexibility, structuring unstructured data better, provide comparison on language-based data as well as provide the solutions needed which is related to the processing of natural language.	
• No indexing needed	

Source: Divya et al. 2020

In their research, Suta et al. (2020), focused that a chatbot can retain discussions with humans by using a standard dialect for an extended period. Chatbot’s application with NLP integration has proven effective in various fields. According to Aprilinda et al. (2022), a chatbot can help students learn English grammar as an independent learning medium. It mentioned that a student can explore the knowledge related to the grammar of tenses they would like to know through conversations with the chatbot. Kavitha and Kavitha (2023) found that chatbots using chat-boats are the conventional means of communicating with computers in natural language. Most chatbots use straightforward text interfaces to converse with their human counterparts, whilst some include speech recognition and text-to-speech capabilities. Their study also found that chatbots can serve effectively in education as well as business organisations.

2.10.1 Implications of Natural Language Processing (NLP) in Current Study

Concisely, Natural Language Processing (NLP) is part of a field in Artificial Intelligence (AI) that allows machine to interpret, process, and provide feedback in common human language. NLP has shown to provide significant possibility to enhance productivity and support unified human-computer interaction.

Despite the advantages listed for the application of NLP such as pattern recognition, cross-model processing, and automation summarisation, the disadvantages of NLP shall be addressed as guidance. The drawbacks of NLP such as high computation costs and limited understanding on the context, disdain, and other visual cues marks the need for further advancement.

The integration of NLP in chatbot proven to be effective in education and business field. The newly proposed chatbot shall be incorporated with NLP as part of its backend process with goals of language modelling which is used to learn the sequences of words used in the interaction with chatbot and trained the system to interact with users or customers well in future in banking sectors.

2.11 TEXT MINING

In this research, the conversations between users and chatbots are recorded, which then will be used for text mining. Text mining also known as text data mining, is the process of transferring unstructured text into a structured format to identify meaningful patterns and new insights. Text mining can be used to analyse a vast collection of textual materials to capture key concepts, trends, and hidden relationships. Hassan et al. (2021) also called text mining as text analytics and identify meaningful information from structured and non-structured text. It is an approach to mine substantial amounts of textual data to get meaningful knowledge (Mateen et al. 2021). It is also “transforming unstructured text into structure format" for a meaningful context and new understanding (Text Mining 2020). Miner et al. (2012), also found from their study, unstructured data can be transformed into meaningful structured data that analysis algorithms.

Siddiqui and Ahmad (2018), discussed in their work text mining algorithms targeted to detect hidden patterns, relationships, and trends for business decision-making. Table 2.10 shows the details for each of the text mining algorithms.

Table 2.10 Text mining algorithm

Text Mining Algorithm	**Description**
K-Means Clustering	• Unsupervised machine learning algorithm
Naïve Bayes Classifier	• Used to group objects. Mostly used in Machine Learning (ML).
Support Vector Machine	• A special classifier to train data provided by label (supervised learning).
Decision Tree	• Used to create training model to predict class.
Neural Networks	• A series of algorithm that detect hidden relationships in a set of data through a process like how human brain functions.
Association Rules	• The rule-based automated learning approach is used to find the meaningful relationship between variables in a large database. It aims to identify the powerful condition discovered in the databases.

Source: Siddiqui & Ahmad 2018

There are six (6) steps of process to be involved in text mining, and this is shown in Figure 2.2.

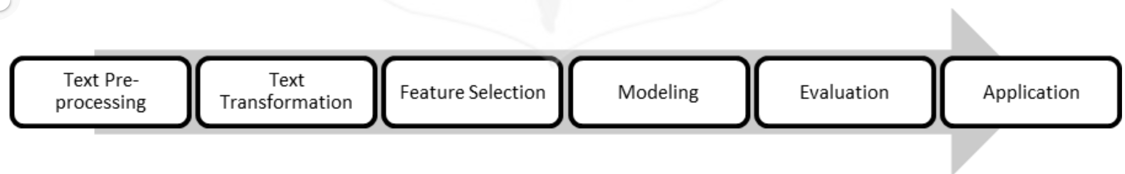


Figure 2.2 Process in text mining

The first step in the process of text mining is text pre-processing. Text processing includes feature extraction and selection. Paramesh and Shreedhara (2019) mentioned that text pre-processing is crucial in the removal of noisy data. The removal of stopwords in text with date, time, and numerical data is assisted by regular expression in pattern matching. During the first step, there are also several sub-steps discovered by Jodha and Dadheech (2019) as follows:

i. Each document is classified into terms called tokenisation.
ii. Removal of stopwords.
iii. Words removal using stopwords list.
iv. Stemming is conducted and completed when the number of features is reduced through the conversion of the derived words into their stems. For example, the root of "cycling" and "cyclist" is "cycle."

Mujilahwati (2016), highlighted in his work that the text pre-processing step was adopted from the Theory of Text Mining. It encompasses the removal of punctuation; case folding which converts all text into small cases; numbers removal either at the beginning or the end of the sentence; removal of white space that corrects the punctuation of the text; removal of words and stop words which include the frequently used and non-important words; stemming that converts the words into its root word; and finally, tokenisation that identifies the words and transforms them into the sequence that are divided based on spaces or special characters. Harith (2022) mentioned various steps of text pre-processing which include tokenisation, lower casing, stop words removal, stemming and lemmatisation.

a. Tokenisation

Tokenisation is the process of replacing sensitive data with unique identification of symbols that retain all the essential information about the data without compromising its security. Tokenisation is also known as word segmentation which involves a process where texts are divided into smaller and meaningful components. The components are categorised into forms such as paragraphs, sentences, phrases, words, and even letters (Geetha et al. 2022). It has also been known that tokenisation works by splicing the sentences individually separated by blank spaces. The unit is called a "token." The tokeniser is required to consider the use of abbreviations, dates, numbers, and other formats as the sentence will lose its meaning when separated. Tokenisation fundamentally will undergo two (2) techniques for text normalisation which are stemming and lemmatisation (Tsakiris et al. 2022). Harshith (2024), showed in his study that the algorithm used for tokenisation involved both the input as well as the output as indicated in Table 2.11.

Table 2.11 Tokenisation

Process	Details
Input	from nltk.tokenize import word_tokenize sentence = “Books are on the table” words = word_tokenize (sentence) print(words)
Output	[‘Books’ , ‘are’ , ‘on’ , ‘the’ , ‘table’]

Source: Harshith 2024

b. Lower casing

Lower casing is done by transforming all words in the sentence into a lower-case form. This step is performed to maintain a consistent flow during NLP processing and text mining (Gupta n.d). For example, if this conversion is not performed, words such as “satisfied” and “Satisfied” will be recognised as two different words which result in duplication.

Example:

tweet[‘Text’] = tweets[‘Text’].str.lower()
tweets.head()

c. Morphological analysis: Stemming

In linguistic morphology and information retrieval fundamentally, stemming is the process of reducing inflected (or sometimes derived words) to their word stem, base or root form which in generally a written word form. Stemming is defined by Divya et al. (2020), as the process to produce a variety of base words. Stemming reduces a base word into its stem word to shorten them for better understanding. Stemming is used for information retrieval systems (such as search engines) as well as to determine the domain vocabularies (domain analysis). Table 2.12 shows examples of root word and the stemming word.

Table 2.12 Example of words used for stemming

Root word	Stemming
Fruit	• Fruits • Fruity
Present	• Presents • Presenting
Design	• Designs • Designing • Designed

In the work on stemming conducted by some researchers, some algorithms represent the Natural Language Toolkit which are Porter Stemming, Snowball Stemming, Lancaster Expression Stemming, and Regular Expression Stemming (Tutorialspoint 2020) which are discussed in more detail below:

i. Porter Stemming Algorithm

Fundamentally, the Porter Stemming Algorithm clarifies every character in each token as either a constant ('c') or vowel ('v') grouping subsegment consonant as 'c' and subsequent vowels as 'v'. The algorithm thus, represents every word token as a specific contribution of consonant and vowel groups.

Porter Stemming Algorithm was developed by Martin Porter in 1980 to substitute common suffixes in English language words (Gawande et al. 2021). This technique is popular for speed and ease which is commonly used for data mining and information retrieval (GeeksforGeeks 2021; Hadni et al. 2012). It is usually used in Information Retrieval Environments (known as IR Environments) for better recall and fetching of search queries and documents represented as vectors of words and the words have the same base word with the same meaning (Tutorialspoint 2020). This technique is said to be equipped with a simple approach for conflation that is applicable to a range of languages as it usually produces less rate of error compared to other stemmer (Willet 2006; Lennon et al. 1981; Ben & Karaa 2013).

To date, there are sixty (60) rules for porter stemming and the general rule for removing a suffix for the Porter Stemming method is as shown below (Porter 2006):

<condition><suffix> →<new suffix>

ii. Snowball Stemming Algorithm

Snowball Stemming is also known as Porter 2 Stemming algorithm, a technique that was derived as an improvised technique from the first Porter stemmer in terms of logic and speed. It is also known as "English Stemmer" (Tutorialspoint 2020). This approach can support fifteen (15) different languages instead of just the English language.

iii. Lancaster Expression Stemming Algorithm

Lancaster Stemmer consists of a set of rules where each rule specifies either deletion or replacement of an ending. It is an algorithm that was developed by a researcher at Lancaster University in the United Kingdom. It is also known as Paice and Husk Stemmer based on a repeated algorithm, developed by Chris D Paice in 1990 (Sirsat et al. 2013; Köhler et al. 2006). There are 120 rules consisting of a table in a rule file. Each rule indicates the removal or replacement of an ending (Rasmi et al. 2021). Boltayevich et al. (2023) studied on the morphological analysis process which looks at methods for creating morphological analyzer and morphological generator, and ways of using the Natural Language Toolkit (NLTK) package in Python. However, Lancaster Stemming is still the most aggressive stemming algorithm and has an edge over many other stemming techniques as it offers the functionality to add customisable in the algorithm when implemented using NLTK package.

iv. Regular Expression Stemming Algorithm

The Regular Expression Stemming Algorithm also known as Regexp Stemmer, is one that utilises a single regular expression, to identify any prefix or suffix and remove from words (Divya et al. 2020).

v. Lemmatization

Lemmatization is a text pre-processing technique used widely in natural language processing (NLP) models to break a word down to its root meaning to identify similarities (Akram et al. 2019). It is also known as the action of mapping a token to the related dictionary which allows the downstream application to be extracted from the "orthographic and inflectional variation" (Knowles & Mohd Don 2004).

d. Stop word

Stop Words known as "stemming" is a technique which involves the removal of structural and unstructured ending words from a sentence or phrase (Shukla & Dixit 2015; Morente-Molinera et al. 2018). In NLP as well as in text mining, stop-words are used to eliminate unimportant words, allowing applications to focus on important words only.

Stop word is also involved in feature extraction. This phase selects a subgroup of main words that represent the text documents. The extraction happens in two steps: document representation and weighting of the feature vector elements. The first step is the document representation. It involves the conversion of documents into vector features. There are several techniques for feature extraction, such as the Named Entity Recognition (NER), Bag-of-words Model (BoW), and TF-IDF as discussed by Tabassum and Patil (2020). NER locates named entities like the name of the individual, organizations, as well as location. NER is also used to differentiate pieces of nouns and obtain clearer concepts of names, countries, and organization (Tabassum and Patil 2020). Bag-of-words is a technique that enables feature extraction from text which is then passed to a classifier. TF-IDF known as Term Frequency and Inverse Document Frequency, is used to correct the words that appear frequently but with less value in any document. TF of Term Frequency detects the frequency of a word that appears in a document without considering the size of the document. Xu et al. (2018) added that the actions provided fair accuracy to both small and huge datasets. While Inverse Document Frequency (IDF) works the opposite way which scales down the word with less significance by calculating the log of ratio from total documents to the total frequency of the word in the document (Tabassum & Patil 2020). The second step is the weighting of the feature vector elements.

The phase for text mining continues with feature selection which is selecting the subset of features from the original documents. It is a procedure that identifies and removes any unnecessary and unrelated characteristics from the feature list which then elevates the accuracy of the sentiment classification (Wankhade et al. 2022). WordNet is a large lexical database of English. SentiWordNet is an extended approach from WordNet that is designed specifically for Sentiment Analysis (Baccianella et al. 2010). Each of the words is assigned to either a positive or negative score.

2.11.1 Implication of Text Mining to Current Study

Concisely, text mining plays significant roles in obtaining meaningful understanding from mass volume of unstructured textual data. The impact varies in multiple areas such as business, healthcare, education as well as research. This was found to be useful to assist any organization to discover concealed patterns, identify the trends, and make decision based on the data obtained.

Text mining provided positive impacts to various industries which can be seen to improve decision-making, automation and task efficiency, improved customer experience, promote risk management and compliance. Text mining helps to identify trends, users' or customers' sentiments, or public view which then notified for better decision-making on marketing, product development, and business strategies. This technique also helps businesses to analyse their users or customers' sentiment such as feedback which enable them to understand and improve the users' or customers' satisfaction. Chatbot which is an AI-driven conversational agent, provides instant feedback through automation and increase efficiency of delivering information to users or customers.

Nonetheless, there is also needed to acknowledge the downside of text mining application such as data privacy concern and uses of distinction of language such as sarcasm and context. By acknowledge these challenges would assist in the advancement of algorithms, techniques of data processing, and ethical aspects.

The newly proposed chatbot makes use of text mining techniques in term of sentiment analysis which the interaction between users or customer were recorded and categorize into three (3) groups which are positive, neutral, and negative. These feedback helps to improve the prototype to meet users' or customers' expectations while using the system.

2.12 ENERGY SAVING AND ENERGY EFFICIENCY CHATBOTS

Energy savings and energy efficiency are features that are so crucial at this time of energy transition and sustainability era. Systems and applications designed and developed must take into consideration these features to help the nation meet the zero-carbon emission by 2050 as promised in the Paris Agreement, COP26 (Glasgow) and COP28 (Dubai). It is stated by Rolnick et al. (2019), that AI technology contributed to global warming issue which become part of initiative arises for looking on sustainable solution in the industry.

Current and future chatbots must ensure that chatbots are energy saving as well as energy efficient to minimize energy consumption, making it more environmentally friendly and reducing the overall impact on device batteries or server resources. Energy efficiency and energy saving is particularly crucial for mobile devices and applications where prolonged usage of resource-intensive processes can drain battery life very quickly (Gamage et al. 2022; Midda 2023).

Research carried out by Singh (2025) on energy consumption of AI powered chatbot, highlighted that generative chatbot used highest energy consumption during both training (150kWh) and inference (5 kWh) sessions. The generative chatbot said to only emit 73.625kg of CO^2 that indicated the most impactful model compared to rule-based chatbot.

An interesting study was conducted by Matheraarachichi et al. (2024) on designing and developing generative AI chatbots for optimised AI chatbots with natural language. Understanding and generating capabilities for the complex information needs of energy IoT infrastructure and net-zero emission. The approach they used for the optimised generative AI chatbot composed of six (6) core modules namely: Intent

Classifier, Knowledge Extractor, Database Retriever, Cached Hierarchical Vector Storage, Secure Prompting, and Conversational Interface with language generator. Based on the experiment they conducted; it was found that the approach had positively assisted human operators in analysing and identifying decision opportunities in the energy grids in a more simplified way for optimising grid operation and net-zero carbon emission efficiency.

2.12.1 Implications of Energy Saving and Energy Efficiency in Chabot on Current Study

The energy consumption for chatbot is related to factors such as size of the application, the performance of hardware used, as well as energy resource used to supply the data center. As stated by Patterson et al. (2021), for generative models which involve widespread large training datasets have the tendency to use more energy compared to simpler models. The used of cloud-based solution can helps to allocate resources well which then results as energy savings strategies better than on-premises arrangements (Khan et al. 2020).

The newly proposed chatbot used middle range datasets which consists of information related to banking services such as loans application, credit application, as well as general information on other banking services. In future, the sizing may increase based on training capacity as well as demand from end users and management team.

2.13 VISUAL ELEMENTS IN CHATBOTS

The integration of virtual technology in chatbots too has become increasingly popular as the visual elements can complement the text-based interactions. This means that instead of relying solely on the text responses, chatbots can use images, graphics or videos to provide information, instructions or engage users and customers in a more interactive and meaningful manner. It can help dispel understanding generally. This improves users or customers comprehension and engagement, making the conversations between the chatbot and humans more intuitive and user-friendly.

Visual chatbots provide a dynamic and engaging user or customers experience with personalised interactions, natural language processing, context awareness to omnichannel integration. Voltolina and Neti (2021) found that creating visual conversational interfaces (VCI) paves the way for modern design challenges. Proper visual cues in chatbots can captivate the attention and engage users' or customers' experience throughout the conversational sessions. By using visual language, the expert developer can specify intents, actions, entities and parameters on which the chatbot can be created.

A study conducted by Goktas and Agildere (2024) recently highlighted the promising potential of large language models such as Open AI's Chat Generative, Pretrained Transformer (ChatGPT) to serve as a supportive tool for radiologic clinical decision making. Its significant improvement was particularly in the visual and image service provided for screening cancers and pain. Such cases and continuous dialogue should continue to show how visuals in chatbots can be beneficial to domains, such as healthcare, education and banking.

Among challenges lies under the application of Natural Language Interface (NLI) is the high users or customers' expectations as they hope to communicate with the system closest to the way they interact with human being (Kavaz et al. 2023).

2.13.1 Implications of Visual in Chatbots to Current Study

The application of visual element in chatbots, specifically in the field such as healthcare, banking and education, shows significant improvement in service delivery and decision making. This advancement of exploration and development of chatbot in visual chatbot technology's view to pave the future AI-oriented communication, resulted in dynamic, information, and user-oriented applications.

The impact brought by the integration between virtual technology and visual element existed in chatbot can be seen in user and customer experience as it is elevated via images and graphics presented in the new proposed chatbot's application. Visual element applied into the newly proposed chatbot comprises the use of visible avatar as persona for interaction, clear use of buttons to represent user's decision and other elements that calculated to help to elevate users' or customers' experience while using the newly proposed chatbot.

2.14 INFORMATION AND INFORMATION-SEEKING BEHAVIOUR

Zhang et al. (2020) mentioned there is research conducted to determine the effectiveness of chatbots in encouraging positive behaviour among users. Thus, this section will focus on behaviour specifically, information-seeking behaviour shows by the users while engaging with chatbots. Information behaviour is defined by Fisher et al. (2005) as "how people need, seek, manage, give, and use information in a different context" whilst information-seeking behaviour is part of human's approach to obtaining information through various resources namely the web browser, newspaper, library, asking a question to peers and consulting the experts (Arowosaye & Bakare 2022).In this research, human information behaviour has become the motivational factor to search for information using the new chatbot developed.

Fidel (2012) mentioned that human information behaviour (HIB) has contributed to "developing conceptual constructs that can guide research, validate the constructs to use in various fields of study in HIB research and discovers the elements that required to be included in HIB research. Fidel (2012) also defined the information interaction of humans as "a complex and multi-layered event that is affected by issues such as: cognitive and affective states; organisational, cultural environment, social aspects; technology, nature, and structure of the source of information". It also stated that there are three (3) focus lines of research in information-seeking behaviour (ISB) as psychological lens, social lens, and in-context lens.

The psychological lens focuses on the study of cognitive and affective aspects of ISB. This branch of ISB, targets the general statements of people looking for information. The social lens specialises in the study of social, cultural, organisational, as well as political aspects of ISB. It focuses on scrutinising the pattern of information-seeking shown by people in specific groups, social constructs, or organizational levels. Lastly, the in-context lens aims to focus on "the actors and their context" which connects the aspects that relate to the actor while searching for information.

HIB which is referred to as the 'information-seeking models' evolution by founded by Tom Wilson. In the model in Figure 2.3, shows Wilson's general model on HIB particularly on information seeking behaviour (ISB). The first four (4) elements, which is context of information need, activating mechanism, intervening variables, and activating mechanism represent the information-seeking behaviour. The "activating mechanism" symbolises the decision to take action to please an information need while the second "activating mechanism" represents the decision to search resources for information (Wilson 1997). Wilson (2005) claims that this model contributed to all research in ISB by illustrating "the attention of the researcher to the totality of information behaviour and showing how a specific piece of research may contribute to an understanding of the whole".

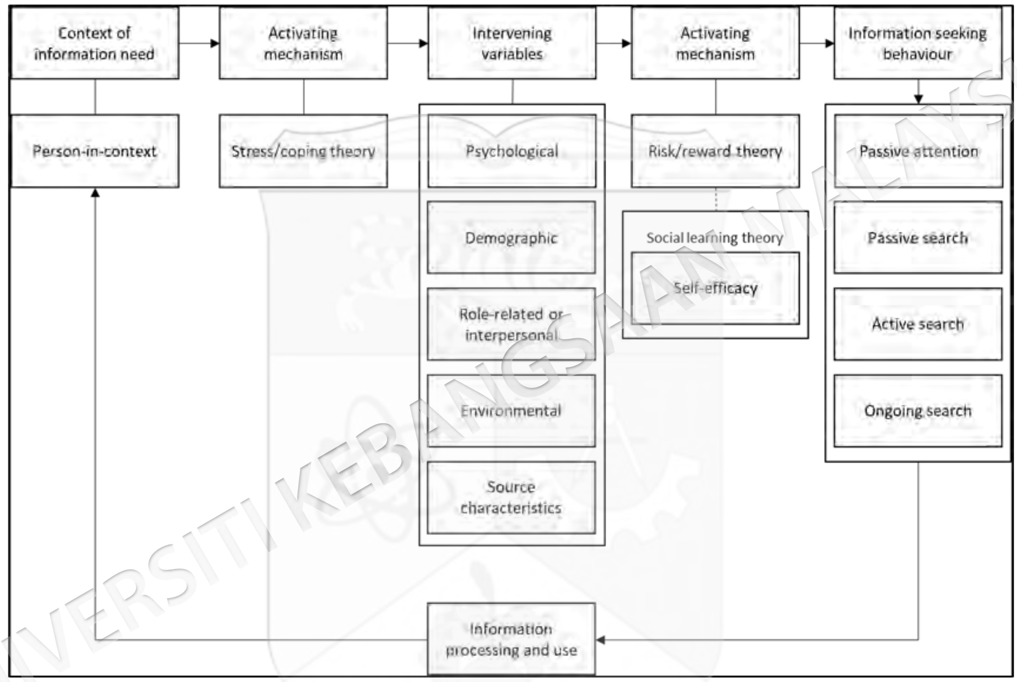


Figure 2.3 Wilson's General Model of Information Seeking Behaviour (ISB)

Source: Wilson 2005

Robert Taylor (1968) introduces the model of prenegotiation decisions by the inquirer which explains in the context of a library search process as described in Figure 2.4. This model highlighted the activities involved before deciding the next action by interacting with three (3) decision points. An interview was conducted among special librarians and information specialists based on several criteria. First, both groups focused on fundamental questions. Secondly, the inquiries presented originated from highly motivated and critical-thinking people. Thirdly, to find material, a librarian must understand and negotiate the nature of the question presented. It is an open-ended interview designed to tackle three (3) questions; a) the category of information the librarian intends to obtain from the inquirer, b) the responsibility of system file organisation in the negotiation process, and lastly c) types of answers or responses to be accepted by the inquirers and what does the effect shows after the negotiation process. There are three (3) decision points labelled A, B, and C. At decision point A, the inquirer will choose whether to ask their colleague or do self-searching through literature or any information centre available. The second decision or point B allows the inquirer to search for answers either from personal files or the library. The decision made would be impacted by previous experience, environment, and ease of access (Taylor 1968). At decision point C, the inquirer's search for information in the library will have two (2) options: whether to ask the librarian through negotiation or to help him/herself using a searching strategy. Taylor (1968) also presented the model of information seeking which outlined five (5) important aspects needed to understand any request which is a) the purpose of the subject, b) the objective and motivation, c) the personal characteristics of the requested, d) the relationship of request details to file organisation, and e) the expected or acceptable answer.

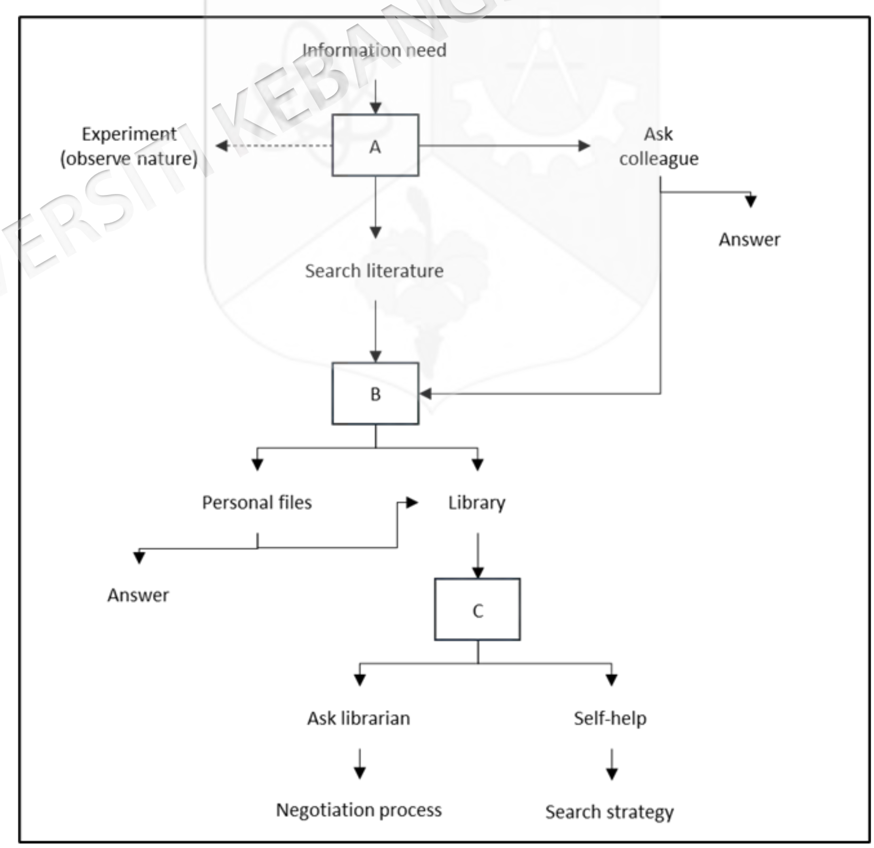


Figure 2.4 Model of Prenegotiation Decisions by the Inquirer

Source: Taylor 1968

Krikelas (1983) developed a model like Taylor's which is known as the model of information seeking. The model shown in Figure 2.5 represents general information seeking which focuses on libraries and the education environment which provides the benchmark for future theories and models in ISB. The models aim to find out the process when a person requires information and observe the elements of information process (Sawant 2015). There are four (4) steps listed for information seeking, a) identifying a need, b) searching, c) finding information, and d) utilising the information. Information seeking is divided into two (2) requirements which are immediate and deferred which determine the method of obtaining the information.

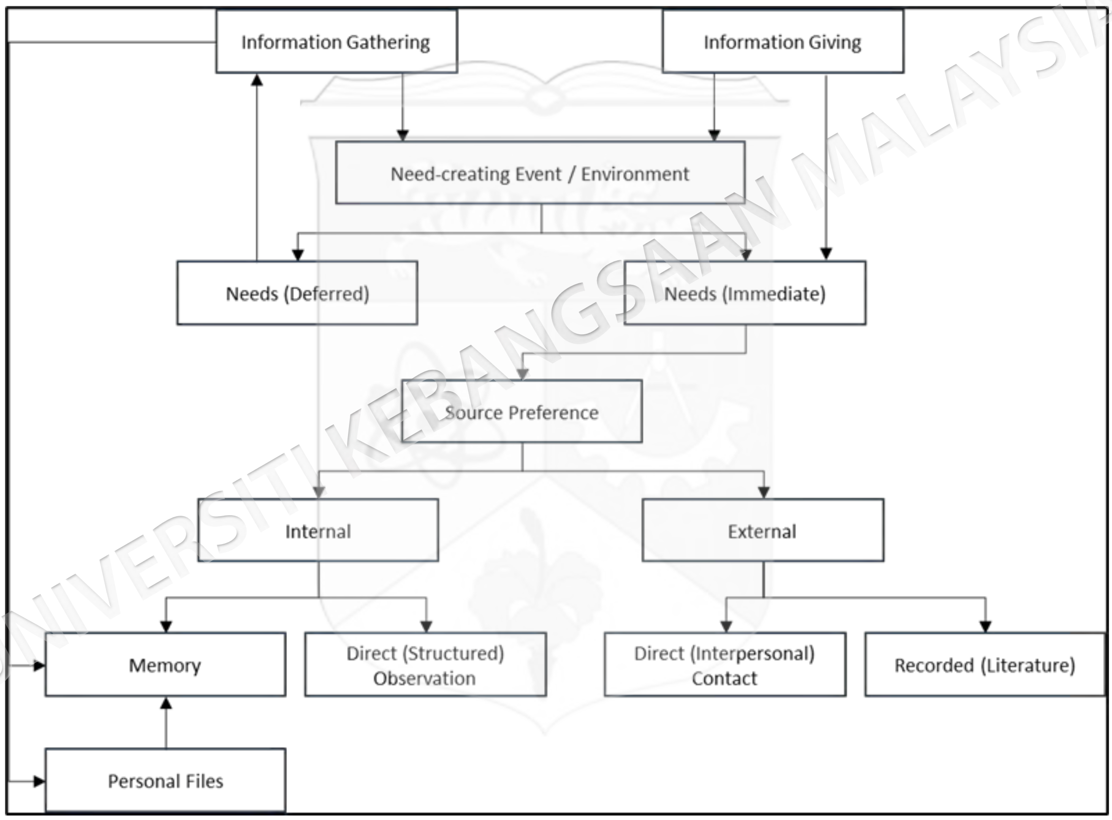


Figure 2.5 Model of Information Seeking

Source: Krikelas 1983

2.1.4.1 Implications of Information-Seeking Behaviour to Current Study

The influence of information-seeking behaviour model in this study reflected in the process of design and development as well as testing of the new proposed chatbot. Adapting the model of information seeking of professional by Leckie et. al (1996) which focus on explore information seeking in three fields: engineering, healthcare, and law. Based on the literature findings, the seeking process comprises of three (3) dynamic elements: sources of information, awareness of information (acknowledgment on source of information and the process of searching), and the result of the seeking process. Figure 2.6 points the elements stated in the model of information seeking for professional against the model adaptation towards the current study.

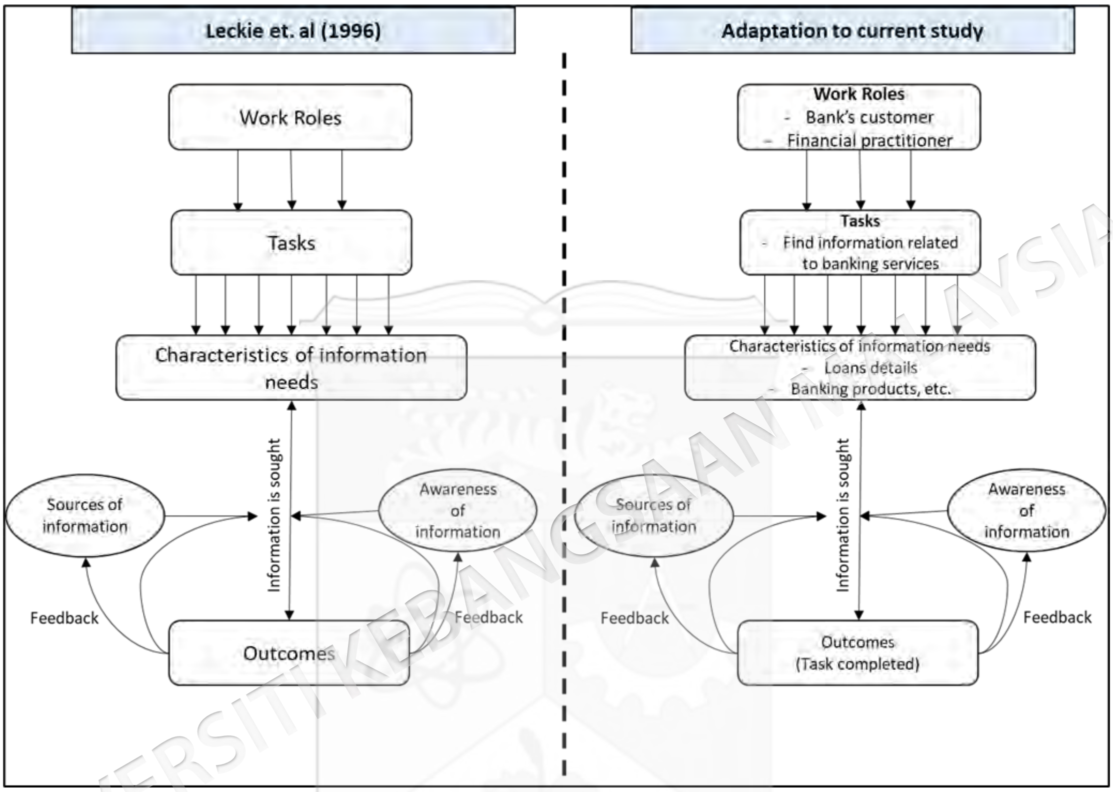


Figure 2.6 Leckie et. al (1996) Model of Information Seeking of Professionals vs. Adaptation to current study

Source: Leckie et. al 1996

Based on Figure 2.6, the process of information seeking started with determining the work roles which the actor for the process. In this study, the end-user of bank's customer and financial practitioners considered as the "actors". Each actor assigned to task which to search for information related to banking services via the new proposed chatbot. The information to be searched listed such as types of services provided, and other banking keywords as listed in the testing instrument. The sources of information used in the chatbot knowledge library was taken from the banking institution sources and integrated to train the chatbot. This is supported by the awareness of the information that is sought and later reflected in the outcomes which indicated the task is now completed. The chatbot being improvised after receiving feedback from end-user or customer and new testing was conducted on the improved chatbot.

2.15 SENTIMENT ANALYSIS

As mentioned earlier, Sentiment Analysis or Opinion Mining, is a natural language processing (NLP) technique used to determine whether data or information presented either in text term or otherwise, is positive, negative, or neutral. Fundamentally, it is often conducted on textual data or information (which are increasingly moving towards a more visual based data and information), to assist organisations or business monitor brand and product sentiment or services, from customer feedback and understand customer needs (Lui 2012). The polarity of the sentences either is leaning towards positive or negative which then will be reflected on their satisfaction level at the end of the conversation. Sentiment analysis is an approach to text analysis (though visual analysis are setting its popularity), whereby it aims to determine the opinion or thoughts and partiality of reviews given by customers (Collumb et al. 2014). The sentiment analysis process comprised of steps such as data retrieval, data extraction and selection, data pre-processing, feature extraction, topic detection, and lastly the data mining process (Hippner & Rentzmann 2006).

According to Cambria et al. (2013), there are three (3) types of sentiment analysis approaches namely: knowledge-based, statistical methods, and hybrid approaches. Trends (2020), elaborated on the methods to execute the sentiment analysis. For knowledge-based, the approach consists of the classification of text based on words that reflect the emotion. Statistical sentiment analysis makes use of machine learning algorithms such as latent sentiment analysis and deep learning for a correct sentiment discovery. Lastly, the hybrid approach, which makes use of both knowledge-based and statistical techniques for immaculate sentiment analysis.

Hussein (2018) is of the opinion that there are four (4) distinct levels of sentiment analysis which cover sentences, documents, words, and aspects. A whole sentence is analysed at sentence level and a score are generated for the sentence. At the document level, the sentiment score is retrieved from the analysis of the whole document essay while at the word level, the analysis is done by making use of each word. Lastly, the aspect level provides the polarity contained in the sentence.

a. Document-level sentiment analysis

The document-level sentiment analysis is used to categorise any chapters or pages of a book into the polarity known as positive, negative, or neutral. This approach applies when the whole document is considered as one entity, and the analysis applied on the whole document. However, the outcome generated from this approach in some cases is not relevant (Vinodhini & Chandrasekaran 2012; Jeyapriya & Selvi 2015). Bhatia et al. (2015) mentioned that there are two (2) approaches to utilise this level namely, supervised and unsupervised approaches to classify the document. Saunders (2021) and Ahmad Fauzi et al. (2024), highlighted that there are two (2) significant areas in this sentiment analysis known as cross-domain and cross-language sentiment.

b. Sentence level sentiment analysis

The sentence level sentiment analysis is used to analyse each sentence of interest from any document with a wide range and variety of sentiment related to it (Yang & Cardie 2014). This level of sentiment analysis is related to subjective classification (Rao et al. 2018). The polarity of the sentence is defined separately through similar methodologies as the document level. However, it requires large training data and processing resources (Wankhade et al. 2022; Su et al. 2023).

c. Phrase/Word level sentiment analysis (WCSA)

Phrase/word level sentiment analysis, focuses on assigning the sentiment score based on the word level by considering the effect of the context towards it from other terms. Wilson et al. (2005) and Su et al. (2023) described this method as valuable for review mining as it can indicate the polarity of the expression (positive or negative) in any sentence.

d. Aspect-level sentiment analysis

The aspect-level sentiment analysis focuses on every aspect used in a sentence which then allocates the polarity to each aspect when the sentiment is calculated (Schouten & Frasincar 2015; Lu et al. 2011). Gu et al. (2024) said that this model first integrates aspect-specific features into contextual sentiment information.

Molina (2022) added that there are other types of sentiment analysis which are known as contextual-based and topic-based sentiment analysis.

a. Contextual-based sentiment analysis

Contextual-based sentiment analysis is used to detect cues and heighten the other types of sentiment analysis. It refers to how words can change the meaning using context which varies from positive, negative or neutral.

b. Topic-based sentiment analysis

The topic-based sentiment analysis is based on specific topics that are known as 'ahead of time'. The analysis breaks up a message into bits of specific topics, which then set a sentiment score to each topic. The topic that is picked from the conversation is then determined into positive, negative, and neutral.

El-Din (2016) summarised the sentiment approach as divided into six (6) stages: Begins with reviews, sentiment word analysis, sentiment identification feature, sentiment classification, sentiment polarity, and sentiment score. There are various levels of tasks in sentiment analysis. Vu and Le (2017) identified three (3) distinctive levels: the first task aims to identify the polarity of the sentiment made in each text whether it points towards positive, negative opinion or neutral. The second is on average level, aims to analyse the level of sentiment according to a scale from 1 through 5 which is labelled from most negative to most positive. Lastly, the third and highest level, aims to classify the expressed opinion towards the distinctive features of a given object.

Sentiment analysis is also of varied types, which specialise in the polarity of the emotion or feelings, which range from positive, negative, and neutral (Joshi et al. 2017). Balvir et al. (2021) highlighted the popular types of sentiment analysis which are: fine-grained sentiment analysis, emotion detection analysis, aspect-based sentiment analysis, and multilingual sentiment analysis.

i. Fine-grained sentiment analysis

The fine-grained sentiment analysis is very often applied as a synonym of the aspect-based sentiment analysis, multiclass classification or sentence-level sentiment analysis. This sentiment analysis, analyses sentence according to parts. It can expand its polarity into scale such as from very positive to positive, neutral, negative, and end with very negative. A new approach for fine-grained sentiment analysis have shown through research by Xiao et al. (2022) that it outperforms six (6) common alternative approaches which are the Naïve Bayes Classifier, Support Vector Machines (SVM), Logistic Regression, Recurrent Neural Network (RNNs) and Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNNs) and Transformer-based Models (BERT, GPT, RoBERTa, etc.).

ii. Emotion detection sentiment analysis

Emotion is defined as a conscious mental reaction (such as happiness or fear) which instinctively experienced as a strong feeling expressed towards a specific object or event and is usually attended with psychological and behavioural changes in the body (Balvir et al. 2021). According to Ekman et al. (1972), humans often express six (6) emotions namely, happiness, sadness, fear, disgust, anger, and surprise. Emotion mining techniques are widely used by leading organisations such as Google, Microsoft, IBM, and others (Pozzi et al. 2016) which are mostly used for their product or service enhancement as well as part of competition analysis.

Thus, emotion detection sentiment analysis allows one to go beyond polarity to detect emotion such as joy, anger, fear, sadness, surprise, and revolt. Many emotion detection systems use lexicons (list of words and the emotion they convey) or complex machine learning algorithms (Machova et al. 2023).

iii. Aspect-Based Sentiment Analysis (ABSA)

Aspect-based Sentiment Analysis (ABSA) is a fine grain type of sentiment analysis that identifies and classifies data according to aspects and opinions from a given text (Shakirah & Ali 2022). The goal of ABSA is to evoke the terms mentioned in targeted sentences and make predictions upon the sentiment in the sentence (Pontiki et al. 2016). ABSA is divided into three (3) categories which are Aspect Category Detection (ACD), Opinion Target Extraction (OTE), and Sentiment Polarity (SP). ACD is used to perceive any entities and attributes in the sentence, while OTE acts by mining the linguistic expression from the sentence, and SP is used to categorise objects or units according to the polarity labels (Pontiki et al. 2016). ABSA is gaining increasingly popularity with the surge of digital opinionated text data, and most importantly, due to its ability to mine more detailed and targeted insights. ABSA also benefitted businesses to understand better on their targeted features of the products or services that are favoured or disliked by users or customers.

iv. Multilingual sentiment analysis

Multilingual Sentiment Analysis is an AI-driven process approach of extracting opinions from data based on several languages. This is achieved through native language Machine Learning (ML) models, built individually for different languages (Biswas et al. 2024). This multilingual sentiment analysis is usually used to decipher clients' responses and reviews on social media in their respective native language. It is crucial as some parts of the world do not use English as the lingua franca (McLaren 2022).

Thus generally, sentiment analysis is often applied in many applications which can be reflected in customer view analysis, and analysis of social opinion towards current issues.

2.15.1 Implications of Sentiment Analysis to Current Study

Sentiment analysis plays a part as indicator to shows the opinion of a person which reflected in either positive, negative, or neutral. The visible impact of sentiment analysis in chatbot is leading to personalized and context-aware interaction between chatbot and the users or customers. By engaging users or customers to express their opinion and emotion will help to improve their support, engagement, sales, as well as brand loyalty.

In this study, the sentiment of the users or customer who using the new developed chatbot is measured in two phases. In the first phase, opinions from the users or customers were used as reference for improvement. Comments received in the testing phase is used to improve the features and service deliverable of the chatbot. The chatbot is then tested in the second phase and opinion from users or customers are collected. Based on the sentiment analysis recorded, it showed that there is positive improvement whereby increasing of positive opinion recorded at the end of testing phase.

2.16 SENTIMENT ANALYSIS ALGORITHM

Sentiment analysis classification techniques are separated into two (2) parts which are namely, Machine Learning (ML) based, and Lexicon based approach (Balvir et al. 2021; Damayati & Lhaksmana 2024).

a. Machine learning based

This approach is a supervised technique. Vu and Le (2017) described this technique as one that applies a widely known machine learning algorithm and makes use of linguistics features to train the sentiment-relevant structures of transcripts. It needs two (2) sets of documents which are training and a test set. The training set is used as the automatic classifier to learn the discerning characters of documents, and the test set is used to examine whether the classifier performs well or vice versa (Vohra & Teraiya, 2013). Naïve Bayes (NB), Maximum Entropy (ME), and Support Vector Machines (SVM) are among the well-established machine learning techniques in sentiment analysis (Vohra & Teraiya 2013). There are many most used features in sentiment classification such as term presence and their frequency, part of speech ambiguity, negations, and opinion words & phrases. Semary et al. (2024), studied in the use of machine learning based sentiment analysis using the feature extraction techniques, which is found in performance to be exactly accurate and superior.

Part of speech is a feature used to label each term in the sentence according to position or role in the grammatical contexts such as proverbs, adverbs, adjectives, and adverbs (Turney 2002; Takele et al. 2024). Negations act to reverse any sentiment (Pang & Lee 2008). Opinion words and phrases will determine the poles of sentiment, whether it is positive or negative. Hu et.al. (2004) and Dilmegani (2024), made use of WordNet and sentiwordnet respectively, to regulate if the extracted adjectives have a positive or negative sentiment.

Another widely used method is the N-gram. It is an approach where it can assign probabilities. Setiawan et al. (2024), conducted a study on the optimisation of N-gram feature extraction based on term occurrence for cyberbullying classification. The findings showed that ideal N-gram patterns were able to increase cyberbullying text characterisation and gave more meaning for context to the field. Each of n-Gram used includes unigram, bigram, and trigram used to test several parameters listed as Minimum Data Frequency (min-DF) and Maximum Data frequency (Max-DF) and term length. Figure 2.7 summarise the research workflow by Setiawan et al. (2024).

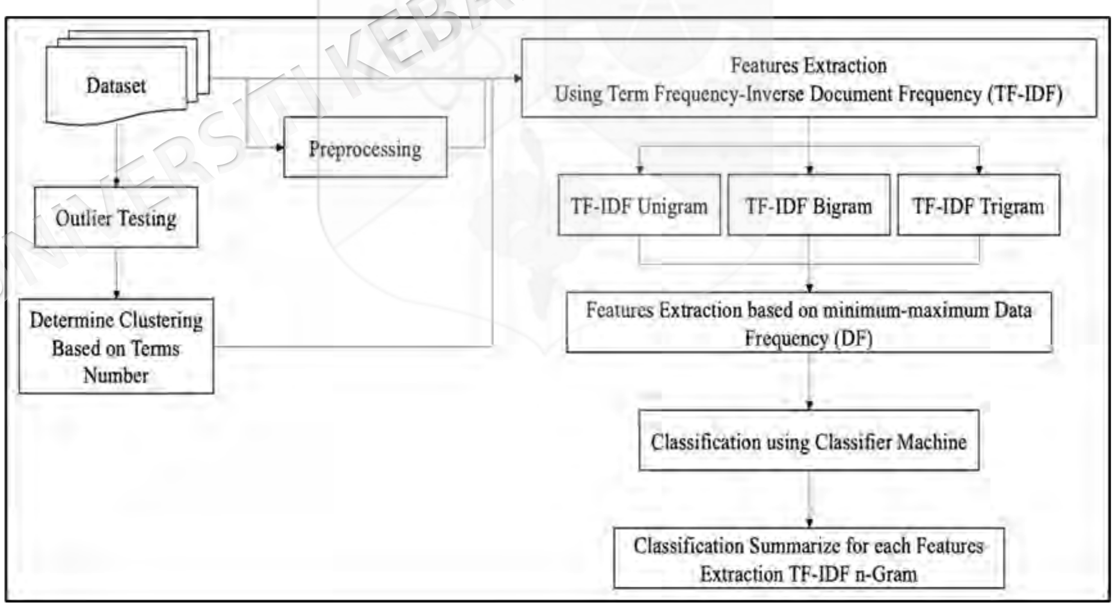


Figure 2.7 Research workflow

Source: Setiawan et al. 2024

b. Rule-based approach

There are three (3) main phases for the Rule-Based Approach which product feature extraction rule learning, opinion sentence extraction, and identification of opinion direction (Yang & Shih 2012; Zahoor & Rohilla 2020).

c. Lexicon based method

Lexicon is defined as "the vocabulary of a person, language, or branch of knowledge" (Lexicon 2019). It is an unsupervised technique which compares text with sentiment lexicons that the sentiment values are classified based on their use (Vohra & Teraiya 2013). Vu and Le (2017), described a lexicon-based method that depends on an assortment of "known and precompiled terms". The Lexicon-based method aims to assign sentiment orientation into groups such as positive and negative sentiment. Meanwhile in another paper Turney (2002), defined the lexicon-based method as an approach which involves orientation calculation for a document from words or phrases listed in the documents. Annett and Kondrak (2008), have listed out the basic steps for a lexicon-based technique which are pre-processed each text, initialising the total text sentiment score, tokenising text, and lastly examining the total text sentiment scores as explained in Table 2.13. Islam et al. (2024), on the other hand worked on the lexicon and deep learning approaches in sentiment analysis on short texts. Their work found that this lexicon based, and deep learning model approach showed a 98% and 95% accuracy, respectively.

Table 2.13 Key steps of Lexicon based sentiment analysis

Key step	**Description**
Pre-processing	• Selecting and remove any noise in the selected document by removing any tags, characters, spelling, and punctuation error as well as any abbreviation used in the text.
Feature Selection	• Extracting the feature appears in the document by using technique such as tagging.
Sentiment Score Calculation	• Set up the initial score, s with 0 (zero). For each sentiment extracted, confirm the availability in the sentiment dictionary. If there is any negative polarity, w then s = s - w, otherwise if there is positive polarity, w then s = s + w
Sentiment Classification	• If the outcome, s, found to be below the threshold mark, thus the document is classified as negative, and vice versa.

Source: Annett & Kondrak 2008

Table 2.14 Comparison between three (3) main sentiment analysis techniques

Approaches	Classification	Advantages	Disadvantages
Machine Learning Approach	Supervised and Unsupervised Learning	• Does not need a dictionary. • Portray high accuracy level for classification	Classifier that trained on the texts for one domain, in most cases does not work for other domain
Rule Based Approach	Supervised and Unsupervised Learning	• Performance accuracy of 91% for review level and 86% at the sentence level • Sentiment classification for sentence scores better than word level	Efficiency and accuracy level is based on the rule defined.
Lexicon Based Approach	Unsupervised Learning	• Labelled data and does not require for procedure of learning.	Needs for vast linguistic resources which is not always available

Source: Devika et al. (2016)

Devika et al. (2016) conducted a comparative study on three (3) approaches that used sentiment analysis techniques. The research considered points for the key factors namely: performance, efficiency, as well accuracy. The results showed that the machine learning approach scored the best result, and this research used this technique. The outcome is stated in Table 2.14.

2.17 CHALLENGES IN SENTIMENT ANALYSIS

Sentiment analysis is often applied in many applications which can be reflected in customer view analysis, and analysis of social opinion towards current issues. However, there are some challenges in sentiment analysis. Pradhan et al. (2016) identified the challenges as firstly a sentiment word can be positive in one situation and can be considered as negative in another situation. The second challenge is people express their opinions differently and slight differences between the two texts did not change the meaning drastically. Other challenges are also addressed in Sentiment Analysis. Pathak et al. (2021) have listed the challenges such as the views on the words which can be in positive form while others perceived the words negatively. Secondly, people tend to express their emotions in an inconsistent way and differently in terms of meaning. Previous text processing found depending on the small gap of variation between two items of text did not change the meaning of the sentence (Selvam & Abirami 2013).

In another research by Wankhade et al. (2022), they have listed a few significant challenges addressed in sentiment analysis which are structured sentiments, semi-structured sentiments, and unstructured sentiments. Structured sentiments focus on formal reviews based on formal issues such as books. It is down among professionals which can express their thoughts and views towards scientific matters. Semi-structured sentiments are between structured and unstructured sentiments which usually consist of short sentences depending on the Pros and Cons listed according to the authors (Birjali et al. 2021; Hussein 2018; Ebrahimi et al. 2017; Mohammad 2017). Lastly, the unstructured sentiments are informal and format-free writing in which the authors are not bound by any conditions (Mukherjee et al. 2013).

In the new chatbot developed, sentiment analysis plays significant role for researcher to identifies the polarity of sentence used while interacting with the chatbot. The words used varies based on the inquiries submitted throughout the conversation. At the end of this research, the outcome from the sentiment analysis was used as supporting indicator for the user's or customer's satisfaction level. The outcome will be considered for future improvement for the services provided by the new chatbot developed in this current study. Liu (2012) mentioned that sentiment analysis can provide valuable inputs on users or customers feedback towards services provided which helps to improve decision making for product development as well as overall service improvement.

2.18 CUSTOMER SATISFACTION INDEX

The performance of any application of software can be measured through the efficiency of tasks completed and the end user's feedback. A chatbot's performance can be measured through the user's satisfaction level. Candello et al. (2017), mentioned that a new conversation system of measurement that might affect the user experience while using a multi-bot system needs to be considered by researchers and the chatbot developer. Customer satisfaction is the key to a successful product being delivered. It provides feedback from users or customers whether it is in positive or negative ways.

The Customer Satisfaction Index (CSI) is a model that represents the voice of the customer (Poliakova 2010). Poliakova (2010), mentioned that the aims of the customer satisfaction index can be categorised into views or opinions, which are helpful for organisations to understand the customers' needs better, provide better enforcement for customer satisfaction based on the features of the products presented, provide better service, allow good visibility as well as to motivate and sense that can enhance the use of product and services provided by the organisations.

Kim et al. (2020) stated that one of the reasonable measurements specifically for chatbot is known as the Customer Satisfaction Score (CSAT). It works to capture the level of user satisfaction towards the performance of the targeted chatbot. Chatbot logic may consist of both dynamic natural language processing or vice versa, and a good category of logic (Trialopa 2022; Gastezzi & Rodriguez 2024).

The Customer Satisfaction Index (CSI) is fundamentally a metric that measures the extent of customer satisfaction on a particular service or product provided by an organisation. The defining nature of characteristics and benefit of the CSI is in a holistic and integrative attribute. This means that CSI does not just measure the overall satisfaction but also specific levels of satisfaction. The CSI would also be able to shed light on which aspects of a business product or service that are meeting, exceeding, or failing the needs of the customer. CSI is a crucial metric to measure consistently overtime, as not doing so, would result to failing to meet customer needs, which would have a negative impact on the organisation's business revenues, profitability, and brand.

The benefit and value of the CSI are several and amongst them are as follows:

i. **Flexibility**: the index is flexible unlike single indicators. It gives opportunity to measure and improve what matters most to the customers, such as product quality, reliability, and trust.
ii. **Identify and prioritise opportunities**: When CSI is conducted consistently, organisations can track the CSI metric, thus allows organisation to prioritise, exploit opportunities to interactively improve customer satisfaction to maintain and build competitive advantage.
iii. **Increase customer retention**: customer satisfaction build loyalty and in turn loyalty compounds results. Loyal customers would mean that they would spend

significantly more and organically attract other customers by spreading positive word of mouth.

iv. **Detect issues and mitigate immediately**: in this hyper-competitive world, the ability to detect the issues early and mitigate immediately is inevitable. Customer would move elsewhere if organisations were too slow to act and mitigate. Thus, customer satisfaction index is beneficial and of value to organisations.

2.18.1 Implications of Customer Satisfaction Index (CSI) to Current Study

The availability to provide service in boundless time, it is shown that less need of human agent and quicker solution can be offered to resolve users' or customers' issues. This event of increasing customer engagement in banking sector which then resulted in higher satisfaction level of users or customers while checking their account balance, transaction history, and transferred fund (Moffitt 2018).

CSI in this research context refers to measuring end users or customer satisfaction level while using the services provided by the new chatbot developed. CSAT Score was used to identify the satisfaction score from both pre and post testing. Based on the feedback received, adjustments are made to resolve issue address during pre-testing phase. Suggestions and comments given used to improvise the initial prototype. Once resolved, post testing was carried out to test the end users or customer's feedback. Based on the score recorded, there are positive engagement changes compare to pre-testing stage.

Sentiment analysis also carried score to determine end users' or customers' satisfaction level. The algorithm provided in new chatbot developed, calculated the sentiment mentioned throughout the conversation. It is recorded that positive changes in use of word in conversation after amendments were made.

2.19 CONCLUSION

The chapter has emphasised on the understanding of the state-of-the-art research on chatbot services that relate to the current study. The literature reviewed in this chapter is divided into seven areas: history, evolution, and benefits of a chatbot, natural language processing, text mining, energy saving and efficiency, information behaviour, sentiment analysis, and customer satisfaction evaluation.

The literature review conducted for the study had reviewed at past research on the areas that are related to this current research and the gaps that exist that need to be addressed. Thus, the areas that were reviewed in detail included the history and state-of-the-art on chatbots and their attributes; sentiment analysis which included topics on information seeking and information seeking behaviour; the types of sentiment analysis approaches using various techniques including Artificial Intelligence (AI) and Machine Learning (ML) and their challenges. The other areas that were reviewed in detail too was on Customer Satisfaction Index (CSI) in general and customer index score particularly for chatbots (CSAT).

The gaps that were discovered throughout the review was the quality attributes that need to be identified and determined to ensure that future chatbots can meet the needs of customers. There is a need to design and develop a chatbot based on the right quality attributes and evaluate the chatbot through sentiment analysis and Customer Satisfaction Index.

CHAPTER III

RESEARCH METHODOLOGY

3.1 INTRODUCTION

This chapter describes the methodology on the entire process of development of the new chatbot. This involves three (3) main phases of design and development. The model used for the design and development of the new chatbot is the standard software life cycle model as indicated in Figure 3.1.

3.2 DEVELOPMENT METHODOLOGY

The design and development methodology model are based on the basic software life cycle model but adapted to the Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle Model (HEI-HC-LC) as can be observed in Figure 3.1.

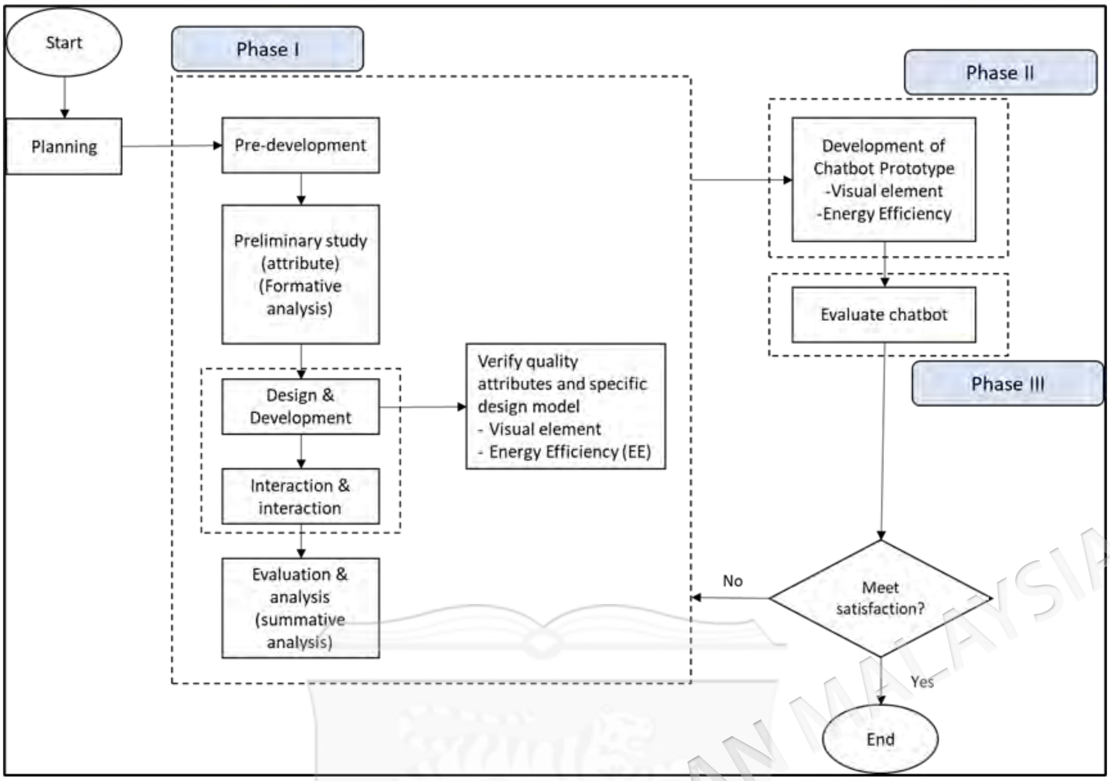


Figure 3.1 Holistic Evolutionary Iterative Humanistic Chabot Development Live Cycle (HEI-HC-LC)

Source: Adapted from Halimah 2018

Based on the diagram, the is a Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle (HEI-HC-LC) approach was undertaken to design and develop the New User Humanity-like chatbot. Details of the phases involved in the model will be discussed in Chapter IV.

3.3 RESEARCH DESIGN FLOW

This research design involved analysis that was conducted both quantitatively and qualitatively to investigate the feedback given by customers upon the usage of a chatbot as well as an improved chatbot. Close-ended questions (items) as well as short-answer questions (items) were designed and constructed in the questionnaire, used as the medium for data collection. All respondents (customers') were required to answer all the questions in the questionnaire which was the instrument used in the study based on the research design as shown in Figure 3.2.

The flow comprised of three (3) phases: pre-development (formative analysis), chatbot development and evaluation (summative analysis). The first phase aims to collect the user's or customer's perception on existing chatbots. Elements involved are the type of chatbot used, suggested features for a decent quality chatbot, evaluation of the existing chatbot and satisfaction level after using the chatbot. The second phase of the study involved the use of a new chatbot developed and assessed based on selected quality attributes through information obtained during the first phase as well as taking into consideration other elements of software development such as functionality and usability. The last phase involved summative analysis based on sentiment analysis of the customers. This phase also involved customer satisfaction index, or the evaluation phase measured that was based on the interaction recorded between the customers and the chatbot.

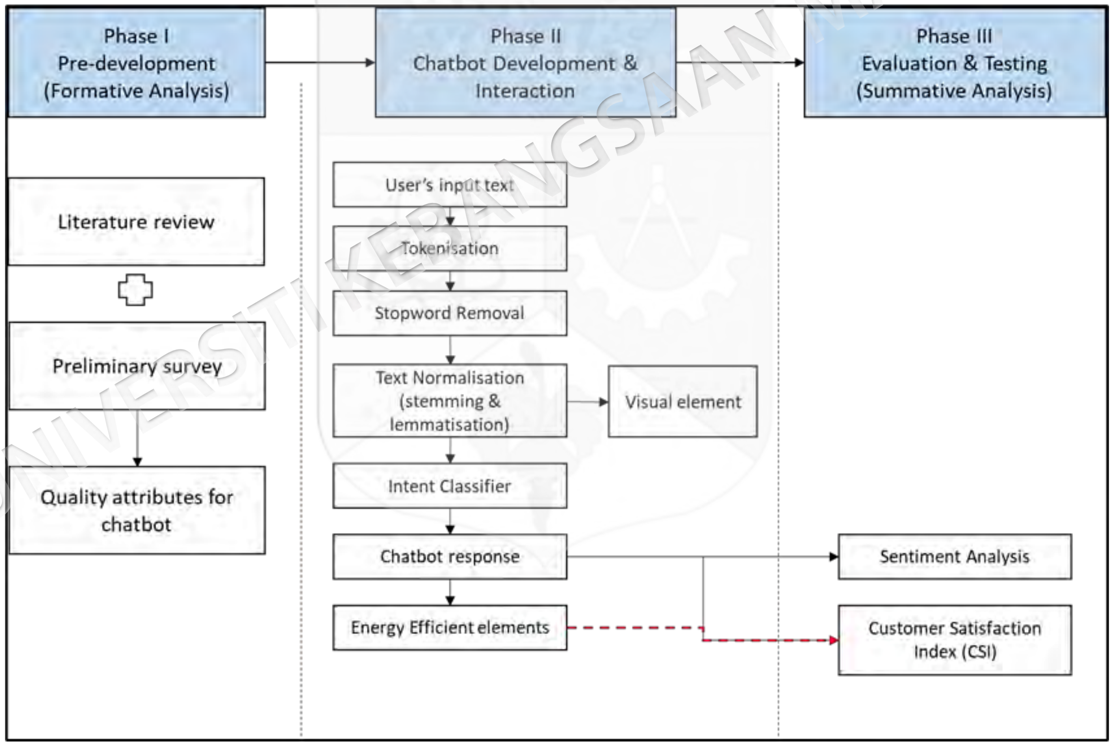


Figure 3.2 Research design flow

3.4 SAMPLE OF RESEARCH

The sample of the research comprised of potential customers of the bank who were selected based on purposive sampling. They were selected based on their being 'customer' of a bank and are familiar with banking operations particularly on retail banking which includes matters on opening an account, whether savings or current; opening a fixed deposit account; applying for personal loans, housing deposit account, car loans; and matters that involved simple corporate banking matters such as shares, corporate shares, and the like. They were also selected randomly based on the different education, race, and ethnic backgrounds.

The samples chosen comprised of thirty-two (32) individuals (customers) for the pilot study. Based on Nielsen (2021) ten (10) individuals are enough to do a usability test of 85% to obtain more than 98% positive results. This study chose to involve triple the number to have a more confirmed positive result (Nielson 2006; Shaba 2024). The User Acceptance Test (UAT) involved ten (10) samples which were selected randomly.

3.5 RELIABILITY OF RESEARCH INSTRUMENT

Rusli and Hasbee (2011) defined reliability as the consistency of any measurement device that is used to produce comparable results from different assessments. Reliability was conducted using Statistical Package for Social Science (SPSS) version 21 for Windows.

3.5.1 Pilot test – Reliability and Validity

All statements in the research instrument were assessed for reliability. Cronbach's alpha value was used to assess the reliability of the instrument design for the research. Section B of the Instrument involved the Evaluation of Chatbot and Section C of the instrument involved Satisfaction Level of Chatbot users or customers assessed for its reliability. The pilot test was conducted, and the Cronbach Alpha value is as shown in Table 3.1.

Table 3.1 Cronbach's Alpha value

Cronbach's Alpha	Internal Consistency
$\alpha \geq 0.9$	Excellent
$0.8 \leq \alpha < 0.9$	Good
$0.7 \leq \alpha < 0.8$	Acceptable
$0.6 \leq \alpha < 0.7$	Questionable
$0.5 \leq \alpha < 0.6$	Poor
$\alpha < 0.5$	Unacceptable

Source: Rusli & Hasbee 2011

a. Results of the reliability and validity tests

The results of the reliability and validity tests are shown in Table 3.2 until Table 3.5.

i. Section B: Chatbot Evaluation Survey: Formative Analysis

Table 3.2 Cronbach's Alpha for chatbot evaluation survey

Reliability Statistics	
Cronbach's Alpha	N of Items
.906	10

The alpha coefficient value for Section B is 0.906 as shown in Table 3.2. Total questions or items evaluated were ten (10) items or questions administered to thirty-two (32) respondents. The value is categorised as to have good consistency value.

ii. Section C: Satisfaction Level of Chatbot's User: Formative Analysis

Table 3.3 Crobach's Alpha for satisfaction level of chatbot's user

Reliability Statistics	
Cronbach's Alpha	N of Items
.722	10

Ten (10) questions or items were evaluated for reliability. The value obtained was 0.722, which was acceptable for its reliability as shown in Table 3.3.

3.5.2 Pre-Survey UAT – Reliability and Validity

The reliability testing was conducted after the pre-survey phase which involved ten (10) samples who were selected randomly regardless of their experience in handling chatbots as well as their background. For the statements in the section, they were selected based on the attributes obtained after achieving the first objective which was to define the best attributes for a chatbot.

a. Section B: Chatbot Evaluation Survey

Table 3.4 Cronbach's Alpha for pre-survey (chatbot evaluation)

Reliability Statistics	
Cronbach's Alpha	N of Items
.955	20

The alpha coefficient value for chatbot evaluation survey was 0.906. Total questions or items evaluated were twenty (20), administered to ten (10) samples. The value was categorised as to have excellent reliability value based on the description shown in Table 3.4.

b. Section C: Satisfaction Level of Chatbot's User

Table 3.5 Crobach's Alpha for satisfaction level of chatbot's user

Reliability Statistics	
Cronbach's Alpha	N of Items
.457	10

Ten (10) questions or items were evaluated for reliability. The value obtained was 0.457, which was unacceptable for its reliability. This value was mostly influenced by the negative response shown by the samples involved in the pre-survey phase towards the chatbot prototype as in Table 3.5.

3.6 RESEARCH INSTRUMENT

The research instrument developed for this research was the questionnaire which consists of three (3) sections. Section A involved the demographic information of the respondents. Section B and Section C involved questions or items on usability and level of satisfaction based on the Likert-scale given. It ranged from strongly disagree to strongly agree as indicated in Table 3.6.

Table 3.6 Likert scale

1	2	3	4	5
Strongly Disagree / *Sangat Tidak Bersetuju*	Disagree / *Tidak Bersetuju*	Neutral / *Neutral*	Agree / *Bersetuju*	Strongly Agree / *Sangat Bersetuju*

3.6.1 Section B: Chatbot Evaluation Survey

a. Pilot study

SECTION B consists of statements that required respondents' responses on the prototype whilst using the chatbot. The statements were adopted from Griol and Callejas (2013), whereby, it was used in the chatbot evaluation. The statements were adapted where appropriate to suit the banking environment. Table 3.7 shows the statements under the section "Chatbot Evaluation Survey".

Table 3.7 Research instrument (chatbot evaluation survey)

No	Statement (s)
1	The system offers enough interactivity.
2	The system is easy to use.
3	It is easy to know what to do each moment.
4	The amount of information displayed on the screen is adequate.
5	The arrangement of information on the screen is logical.
6	The chatbot is helpful.
7	The chatbot is interactive.
8	The chatbot reacts in a consistency way.
9	The Chatbot complements the activities without distracting or interfering with them.
10	The Chatbot provides adequate non-verbal feedback.

b. Pre-survey phase

During the pre-survey phase, a separate set of questions or items were used in the instrument based on the categories of attributes obtained whilst answering the first research objective which were namely, functionality, efficiency, technical satisfaction, humanity, effectiveness, and ethics as can be observed in Table 3.8.

Table 3.8 Research instrument for pre-survey (chatbot evaluation)

No	Category	Statement (s)
1	Functionality	• Able to interpret keywords or instructions accurately • Flexible to interpreting knowledge • Accuracy of the response provided • The information provided is secured
2	Efficiency	• Easy to use • Quick response to questions asked • Can be used at anytime and anywhere • Able to be accessed on any platform
3	Technical Satisfaction	• Able to convey greetings • Able to provide emotional information using right tones and expressions • There is no need for technical skills or proficiency to use the chatbot • The chatbot provides information in a fun, interesting, and innovative way
4	Humanity	• Realness of robot • Able to create enjoyable interaction • Able to maintain the theme of the conversation • Can mimic human personality
5	Effectiveness	• Linguistic accuracy • Have a wide range of knowledge (related to banking)
6	Ethics	• Trained with knowledge culture and ethics of users or customers • Sensitive to user's or customers concerns

3.6.2 Section C: Satisfaction Level of Chatbot's User

SECTION C of the questionnaire consist of statements that required the samples to rate their satisfaction level after using the chatbot. The set of statements were adapted from the System Usability Scale (SUS), a ten-item (10) scale which covered the subjective assessment of usability. This instrument was initially developed as part of the usability engineering program in integrated office system development at Digital Equipment Co Ltd, Reading, United Kingdom (Brooke 1996). The instrument was adapted to meet the need of this research. Table 3.9 shows the statements listed under the section "Satisfaction Level of Users or Customer".

Table 3.9 Research instrument (satisfaction level of user or customer)

No	Statement (s)
1	I think that I would like to use this chatbot frequently
2	I found the chatbot unnecessarily complex
3	I thought the chatbot was easy to use
4	I think that I would need the support of technical person to be able to use this chatbot effectively
5	I found the various functions in this chatbot were well integrated
6	I thought there were too much inconsistency in this chatbot
7	I would imagine most people would learn to use this chatbot very quickly
8	I found the chatbot very cumbersome to use
9	I felt very confident using the chatbot
10	I needed to learn a lot of things before I could get going with this chatbot meaningfully

3.7 DATA ANALYSIS TECHNIQUE OF RESEARCH

After the data was collected during both the phases, data was analysed statistically using the technique of inferential statistics. The research objectives as well as hypotheses were evaluated using statistical analysis methods. A pilot study was conducted with thirty-two (32) respondents. As for the actual usability testing, sixty (60) respondents were selected randomly to participate in the test. Summary of the data analysis method used is as indicated in Table 3.10.

Table 3.10 Statistical analysis method

Item assessed	**Statistical Analysis**
Demographic Information - Respondent Information - Experience of Using Chatbot - Features of a Good Chatbot	Descriptive Analysis
Customer Satisfaction Level	Descriptive Analysis - Collected using questionnaire and text analysis
Sentiment Analysis	Algorithm designed to indicate the polarity of the sentence used (i.e positive, neutral, and negative).

3.8 OVERALL METHODOLOGY

Table 3.11 shows the overall summary of the research methodology adopted in the study. As seen, the methodology adopted is one that is of mixed approach: the standard software development life cycle methodology approach and the triangulation approach of evaluation to ascertain the ability of the chatbot and the user or customer satisfaction level.

Table 3.11 Overall methodology adopted

	Phase I Pre-development (Formative analysis)	Phase II Development Integration & Implementation	Phase III Evaluation & Testing (Summative analysis)
Development Methodology	Evolutionary Iterative Approach to Chatbot Development Life Cycle Methodology		
	• Planning	• Verify quality attributes of new chatbot	• Evaluate and test and CSI
	• Preliminary study • Quality attributes of chatbot	• Pilot Study and development	• Usability Testing
		• Text Analysis • Stemming • Lemmatisation	• Actual study • Sentiment analysis • Customer Satisfaction (CSAT)
Outcomes	RO1 – RO4 RQ1 – RQ4	RO1 – RO2 RQ1- RQ2	RO3 – RO4 RQ1 – RQ4

3.9 CONCLUSION

Thus, this chapter had discussed in detail the development, implementation, and evaluation methodology adopted in this research. The overall methodology involved is the mixed mode approach, adapting the standard system (chatbot) development methodology in this case, the Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle (HEI-HC-LC) model and the triangulation of data mining both the qualitative and quantitative approach to data analysis when ascertaining the usability of the chatbot developed and the user or customer satisfaction Index based on the use of the new chatbot. The methodology adopted is suitable to ensure the enhancement of customer experience and to ensure their satisfaction whilst using the chatbot. Through the sentiment analysis and user interaction with the chatbot, the level of customer satisfaction can be ascertained. In the process both developer as well as the customers benefitted from the experience. A panel of experts were involved in the validation process to ensure that the findings were solid and can be accepted and verified underpinned by the various theories depicted in the philosophical and theoretical approach of the research.

CHAPTER IV

DESIGN AND DEVELOPMENT

4.1 INTRODUCTION

This chapter discusses on the detailed design and development aspects to achieve the appropriate objectives and research questions of the research based on the three (3) phases of the New User Humanity-like Chatbot (NURUL-C) framework namely: the pre-development which includes preliminary study and the formative evaluation; the development, integration, and implementation; and the evaluation and testing which includes summative evaluation.

4.2 CHATBOT FRAMEWORK

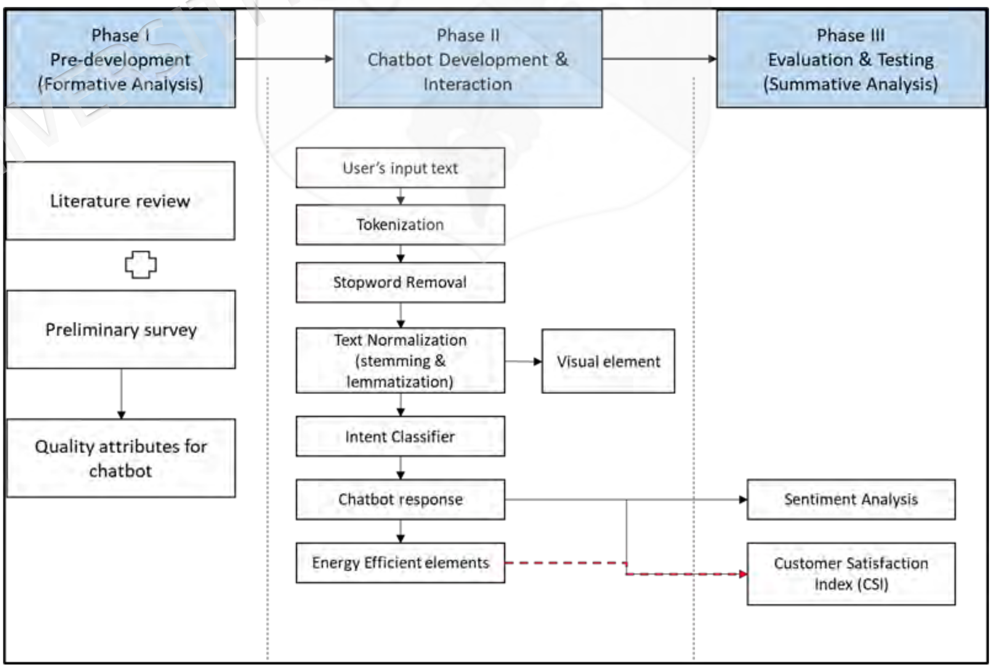


Figure 4.1 Research framework of the New User Humanity-like Chatbot (NURUL-C) (adapted from "Chatbot Engine Process")

Source: Tsakiris et al. 2022

Figure 4.1 shows the NURUL-C Framework based on the three (3) phases mentioned earlier. Details of the activities involved in the three (3) phases based on the Holistic Evolutionary Iterative Humanistic Chabot Development Live Cycle (HEI-HC-LC) as seen in Figure 4.2 until Figure 4.4. Figure 4.1 on the chatbot framework focuses on three (3) phases of development, the pre-development, development of the chatbot, and the evaluation phase or post-development.

4.2.1 Phase I: Pre-Development

The first phase of the HEI-HC-LC is the pre-development phase which focuses on gathering information from the existing chatbot's user using close-ended surveys as medium through data obtained from past research conducted on chatbots. The instrument used was adapted from past research which has undergone reliability and validity tests conducted in this phase. Figure 4.2 shows detail of the activities in this phase.

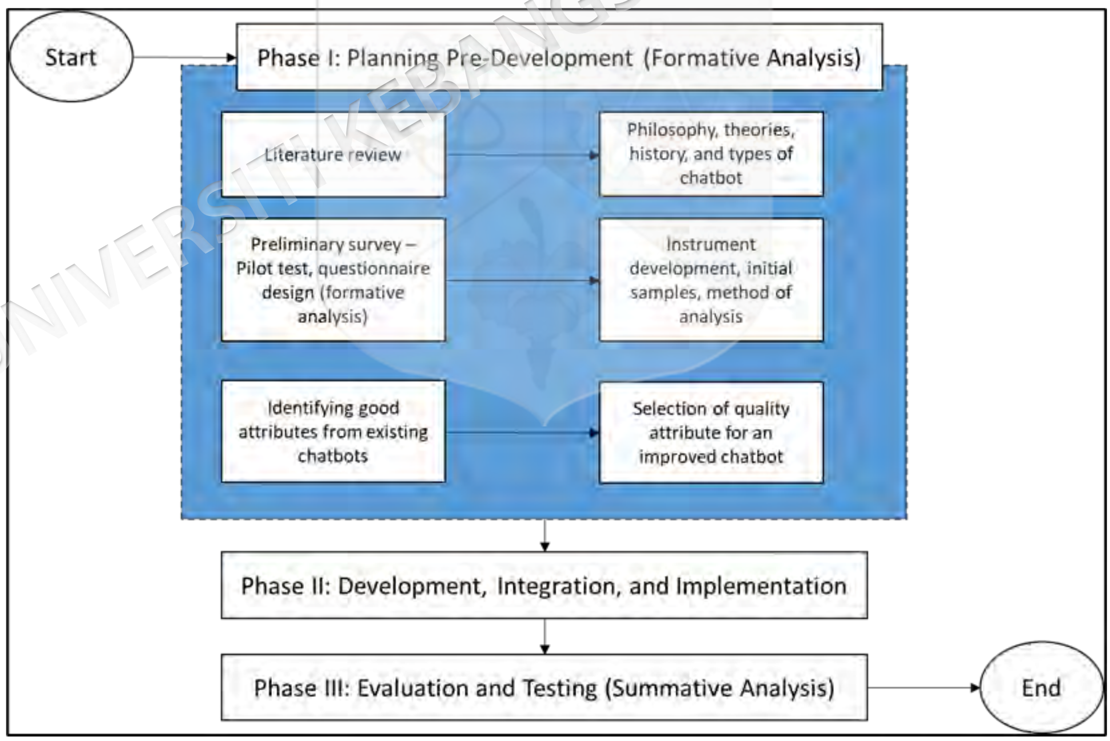


Figure 4.2 Phase I: Planning pre-development (formative analysis) of the HEI-HC-LC

4.2.2 Phase II: Development, Integration and Implementation

The next phase is the development, integration, and implementation phase of HEI-HC-LC based on System Requirement Specification (SRS) of the chatbot. The flow starts with user interaction with the chatbot whereby they key input is the text input. It can be in a statement or questions. Then, it will undergo a tokenization process to chunk the sentences into units. Once done, stopwords were removed and moved to text normalisation. The process continues with the intent classifier shown in the response given by the chatbot. In this phase, there are multiple frameworks integrated at both front-end and back-end as discussed earlier in the Research Methodology used to develop the chatbot. Figure 4.3 shows the detail activities involved in this phase.

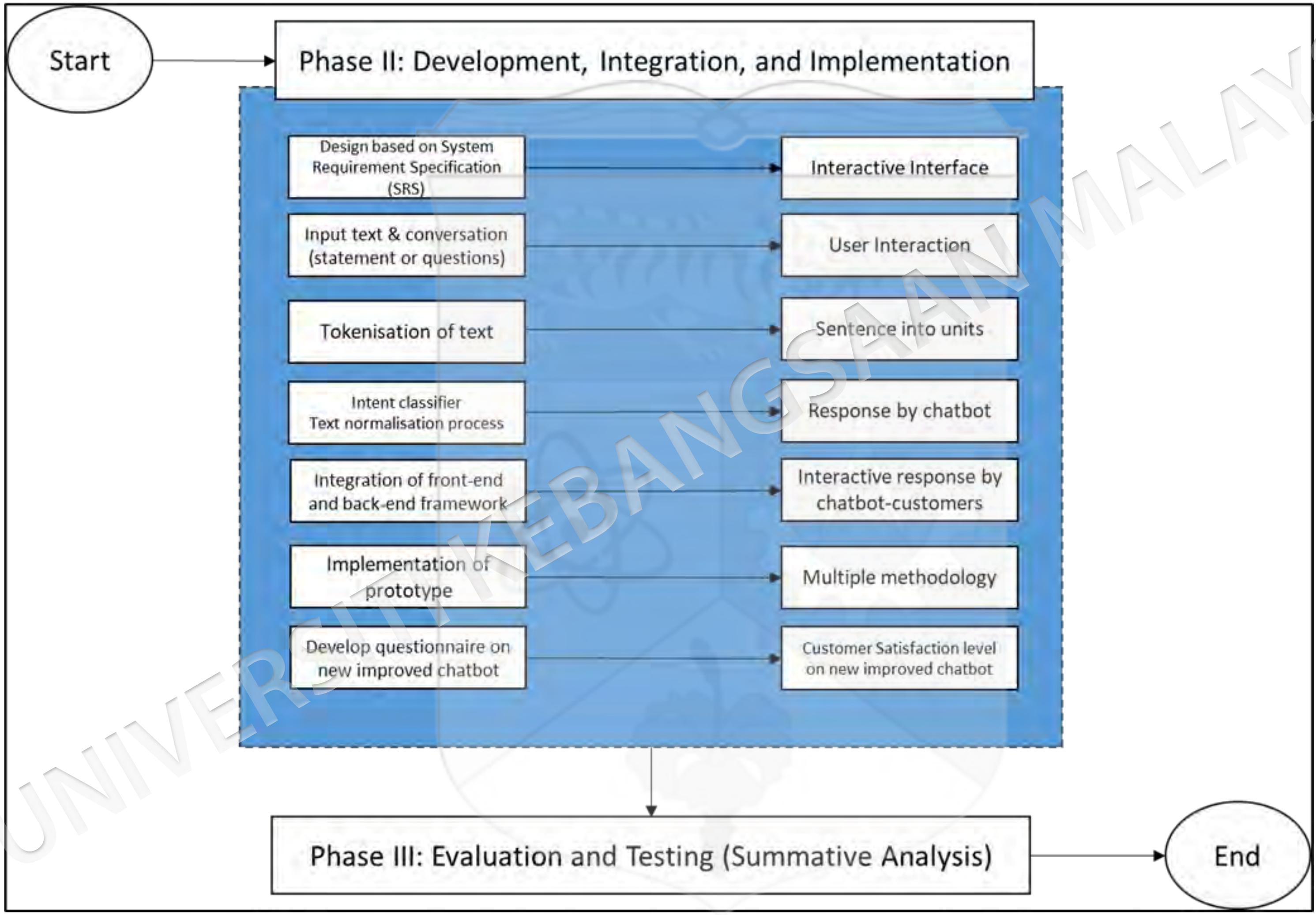


Figure 4.3 Phase II: Development, integration, and implementation of the HEI-HC-LC

4.2.3 Phase III: Evaluation and Testing

The final phase of the HEI-HC-LC is the Evaluation and Testing (Summative Analysis) which focused on the sentiment analysis and the usability test (Summative Analysis) which indicate the level of satisfaction of the users (customers) on the newly built chatbot. Figure 4.4 shows the details of the activities conducted in this phase.

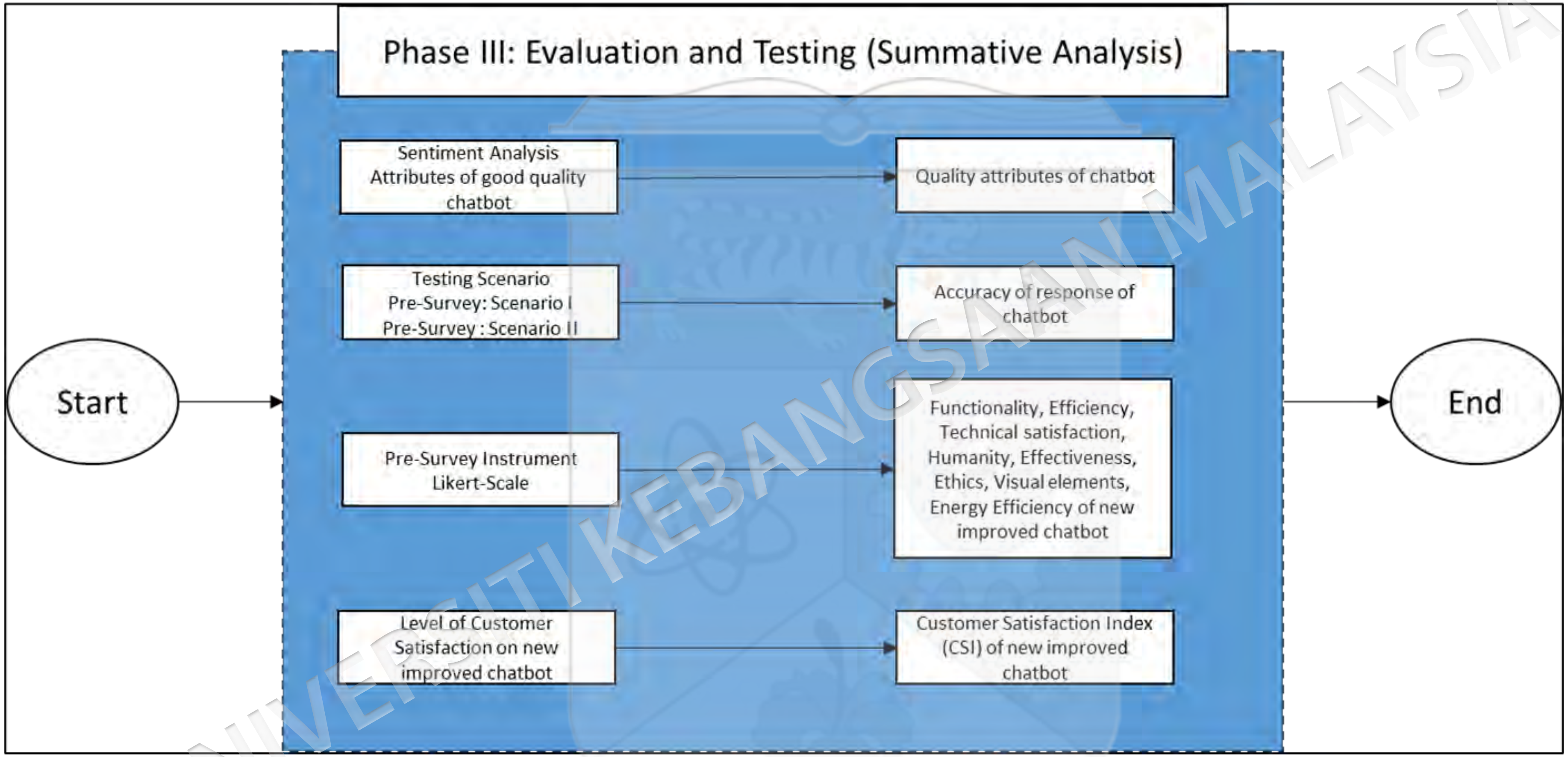


Figure 4.4 Phase III: Evaluation and testing (summative analysis) of the HEI-HC-LC

4.3 CHATBOT INTEGRATION AND FLOW

The outcome of the preliminary stage (Phase I) integrated into the newly developed or improved chatbot called NURUL-C at the second phase. The overall phases consist of pre-development; development, integration, and implementation; and evaluation. The NURUL-C chatbot developed using two-ends frameworks which is front and back end. The front end consisted of three (3) frameworks that work concurrently, which in turn interacted with the API to the back end of the chatbot. After the handshake took place, the correct response sent to the user passing the same API to the front end. The diagram shows the chatbot integration model and framework that explains the interaction flow between the front-end, the API, and the back-end.

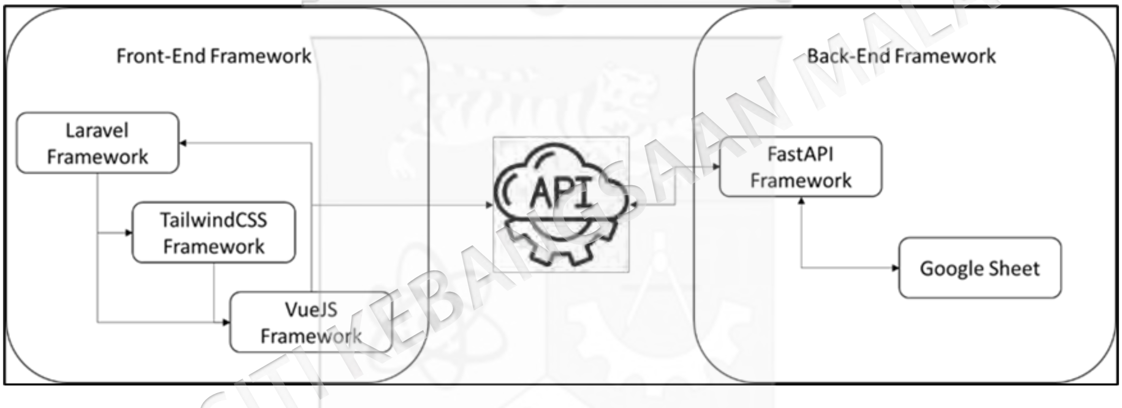


Figure 4.5 Chatbot integration model/framework

Figure 4.5 shows the Chatbot Integration Model / Framework used in the development of the chatbot. The front-end framework consists of; (i) PHP Laravel framework; (ii) Tailwind CSS framework; and (iii) VueJS framework. PHP Laravel framework is a web application framework with meaningful and sophisticated syntax (Laravel 2023). It is also known as one of the frameworks that mostly used by many developers. Ibrahim (2018) claimed that this framework is favourable for complex enterprise-level applications which then can escalate the time and efficiency of any institution's information system. TailwindCSS and VueJS were responsible for building and designing the interface for the chatbot.

The back-end framework consisted of the FastAPI framework and Google Sheet used in this process which functioned as the sentiment analysis reference sheet. FastAPI acts as the high-performance, modern, web framework for APIs construction with

Python 3.7+; the key features listed as fast to code, fewer bugs, easy to design and learn, minimum code duplication with open standards for APIs includes OpenAPI and JSON Schema (FastAPI n.d).

The operation of the chatbot runs was based on the rules from three (3) entities which were the end-user, chatbot, and developer or administrator. The developer provided the necessary insight for the chatbot which then trained throughout the interaction with the end-user. Figure 4.6 summarises the interaction flow between the entities or the context level diagram of the chatbot involved in the process.

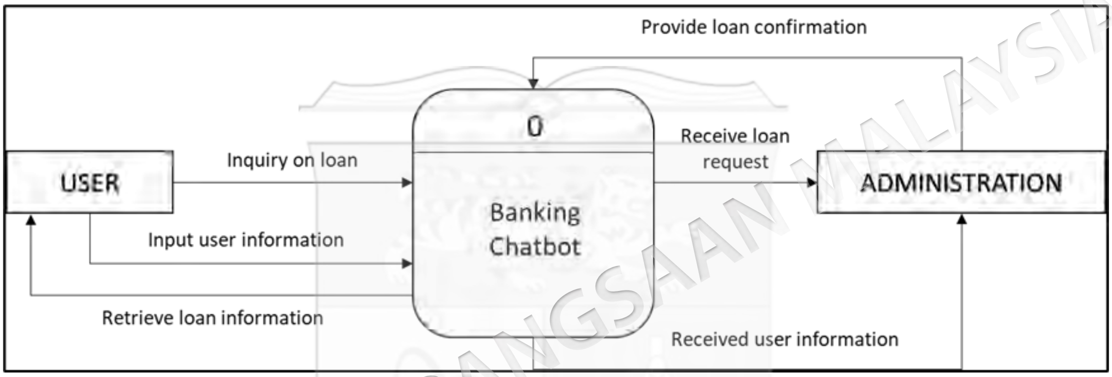


Figure 4.6 Context level diagram of chatbot

4.4 DATA COLLECTION METHOD

Samples of this research were selected both purposively and randomly using the technique of snowball where the respondents will share the survey form among their group. Each data collected throughout the research used were both quantitative and qualitative data to support and answer the research questions. The items or questions prepared in the instrument (questionnaire) were administered to the respondent using Google Forms as platform in two phases: Pre-development and the Development, integration, and implementation phase.

a. Pre-development: Formative analysis through pilot test

At the pre-development phase, where the formative evaluation was conducted through a pilot test, a structured questionnaire was designed based on the existing features of the chatbot ascertained from past research. The main aim of the preliminary pilot study was to determine the quality features of a chatbot that would be determined to build an improved chatbot later. Figure 4.7 shows the pre-development flow: Formative Analysis through Pilot Study.

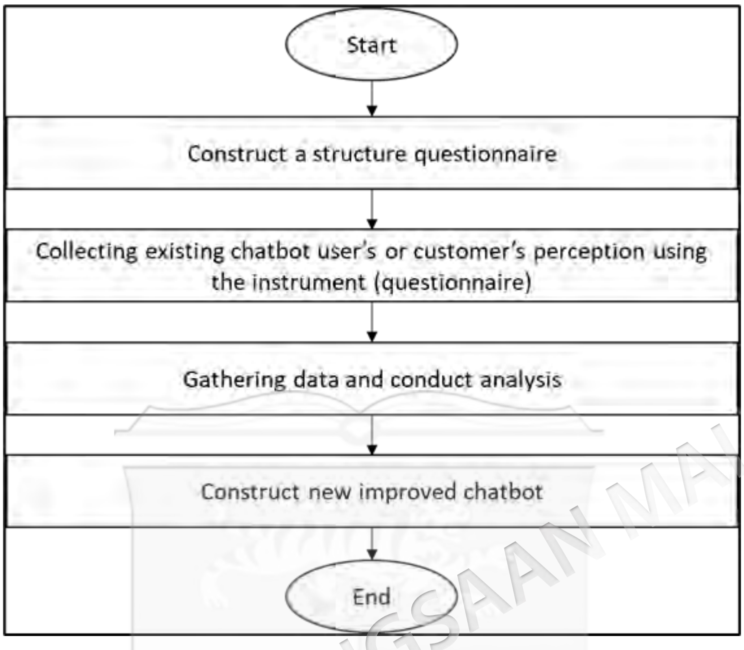


Figure 4.7 Pre-development flow: Formative analysis through pilot study

b. Analysis flow

The result of formative analysis flow was used to reconstruct a structured questionnaire to analyse and ascertain the level of satisfaction users or customers during the development, integration, and implementation as well as the Evaluation and Testing (Summative Analysis) phase. The general chatbot analysis flow that was conducted as is indicated in Figure 4.8.

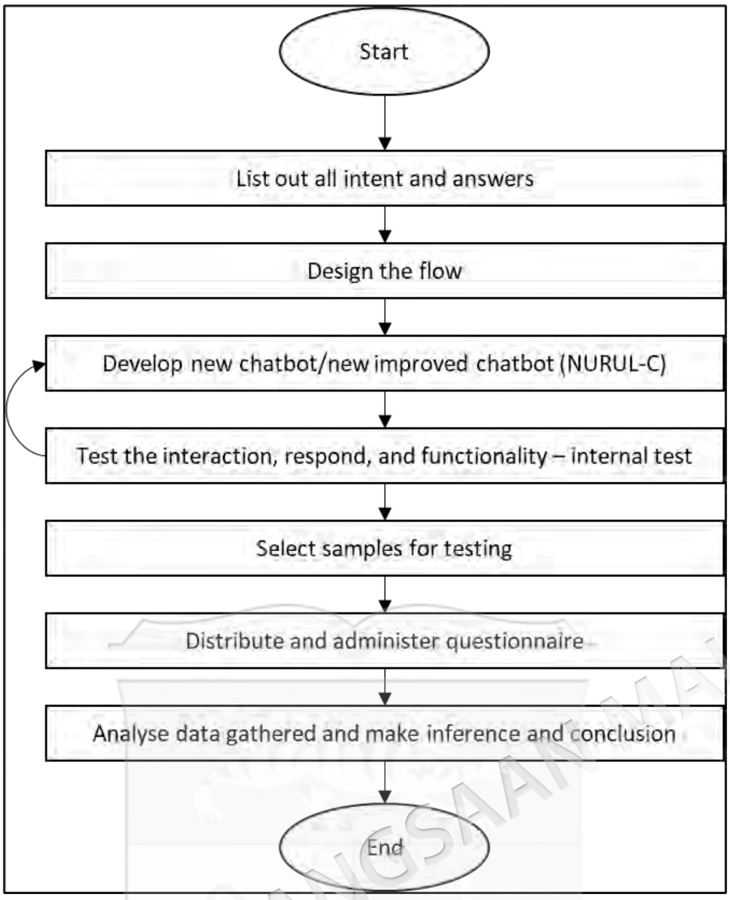


Figure 4.8 Chatbot analysis flow

For chatbot evaluation and testing, there were two (2) phases implemented which was the pre-assessment and post-assessment phase. During pre-assessment, a group of ten (10) samples (users or customers) are selected to participate to evaluate the developed chatbot. Any remarks on the functionality, ease of use of service provided, and issues that arise during the testing were carefully noted. The information and data gathered were later used to improvise the improved chatbot and will be evaluated at the post-evaluation phase (summative analysis). A group of ten (10) samples were again selected to participate in this phase. They were required to assess the chatbot; interact with it, administered and fill up the questionnaire. The sentiment analysis and satisfaction level of the users or customers were done concurrently, while the samples were engaged with the chatbot.

The data obtained from the two-phase questionnaires were analysed using statistical analysis tools and the conversation script that recorded the interaction of the chatbot with the user was used to conduct semantic analysis and the satisfaction level of the users. Once the data was collected from the questionnaire, it was analysed using statistical analysis. As for the semantic analysis, a lexicon-based technique was used whereby the sentence from the user-chatbot conversation was analysed based on the words and expressions used to classify the sentence as positive, negative, or neutral. Semantic analysis is also the medium to determine the customer satisfaction level as well as the data collected from the instrument: the questionnaire.

4.5 DEVELOPMENT OF CHATBOT: INTERFACE ADEQUACY

The development of the chatbot conducted at both the front and back end. The interface adequacy is important aspect of the chatbot and built based on the user or System Requirement Specification (SRS) for the banking domain. The chatbot is accessible through a browser which is on the local device. The chatbot is not in published mode and only be accessed locally.

Two (2) aspects of the SRS that had to catered to before the development of the chatbot began were the hardware and software requirements as observed in Table 4.1.

Table 4.1 Chatbot System Requirement Specification (SRS) hardware and software

System Specification	Requirements
Hardware	• Computer that runs using MacOS, Windows 10 or equivalent • Processor – Intel Core i5 / M1 Apple Chip • Android device / IOS Device
Software	• Programming language: Python, PHP, JavaScript • Scripting Language: HTML, CSS • Framework: Laravel, FastAPI, TailwindCSS, VueJs

Figure 4.10 until Figure 4.12 shows examples of the print screen of the user interface of the chatbot. Figure 4.9 shows the welcoming interface of the chatbot, that is clean and straight to the point- implying in professional approach to the users or customers, customary of a bank. The background colour used is based on the bank’s theme colour. Simple instruction stated at the opening page for users or customers to start the conversation by clicking the “Start Conversation” button.

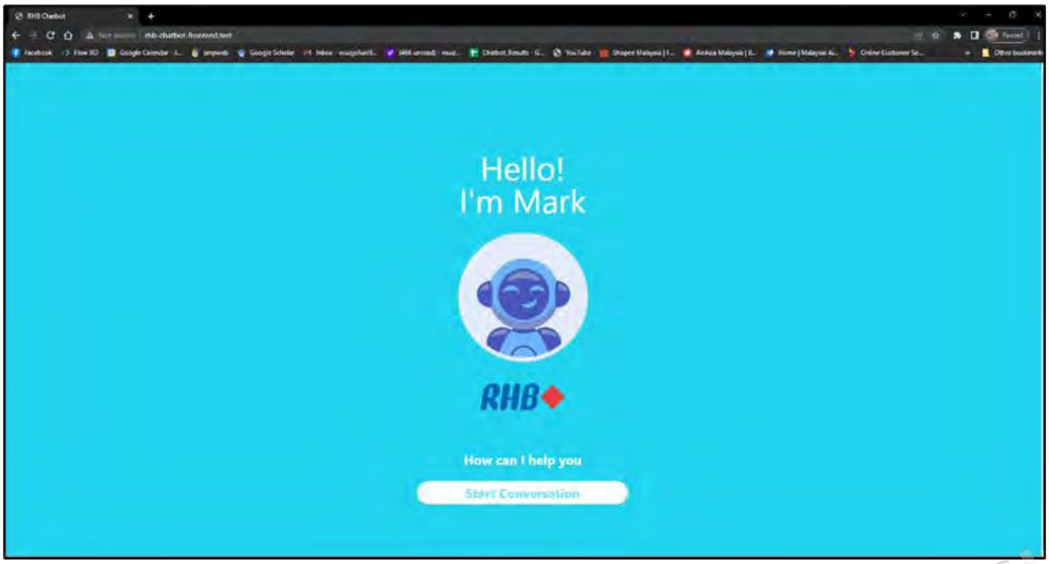


Figure 4.9 Welcoming interface of chatbot

Figure 4.10 shows the intent content interface of the chatbot where users or customers can choose to make interaction with the chatbot. The icons at the top shows name of the new improved chatbot and conversation opening by the chatbot. A clear instruction was included for the user or customers to start the conversation by input any keywords or click on the button provided on the bottom of the page. These buttons allow the users or customers to choose and interact with the chatbot based on the services that are required. The interface is designed as such based on the information seeking behaviour theory of users. Whenever they are seeking for information, they want to ensure that all interactions to be done are visibly seen through the buttons designed as the interface. This eases the users of the chatbot.

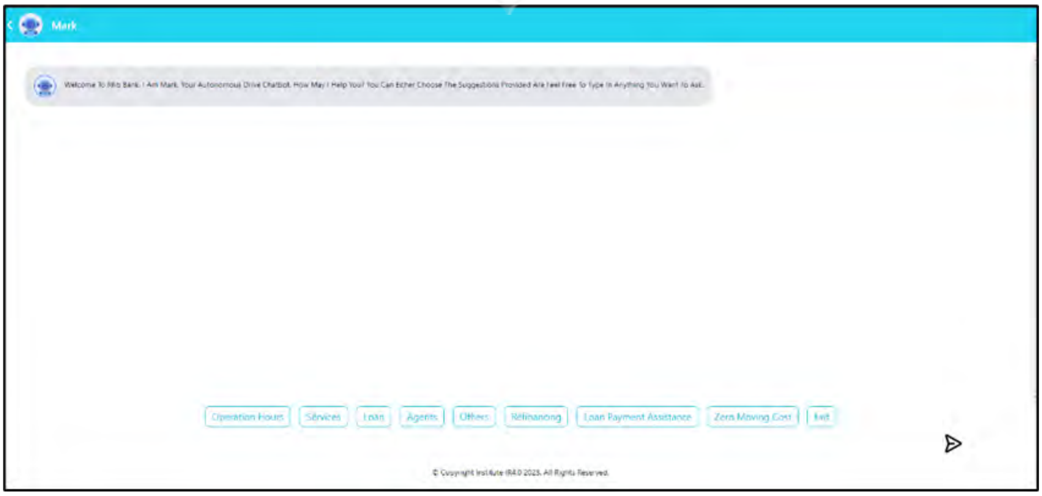


Figure 4.10 Interface for chatbot

Figure 4.11 shows the interaction record between the chatbot and the users. Similarly, the interface is clean, solid and user friendly. The design is clean to depict the straightforwardness, transparency and trusting characteristics of banking world.

Figure 4.11 Interaction record between chatbot and end users or customers

Figure 4.12 shows the mobile interface of the chatbot. Overall, the interface is simple to use, friendly, and clean. The approach is in line with the Information Seeking Behaviour (ISB) theory and customer behaviour theory that would like all interactive buttons to be easily accessible. The greetings by the chatbot also appear to the users or customers as they feel at ease when interacting with the chatbot, as this gives a feeling of realism when interacting with the 'robot'. The element of transparency and trust too becomes more realistic.

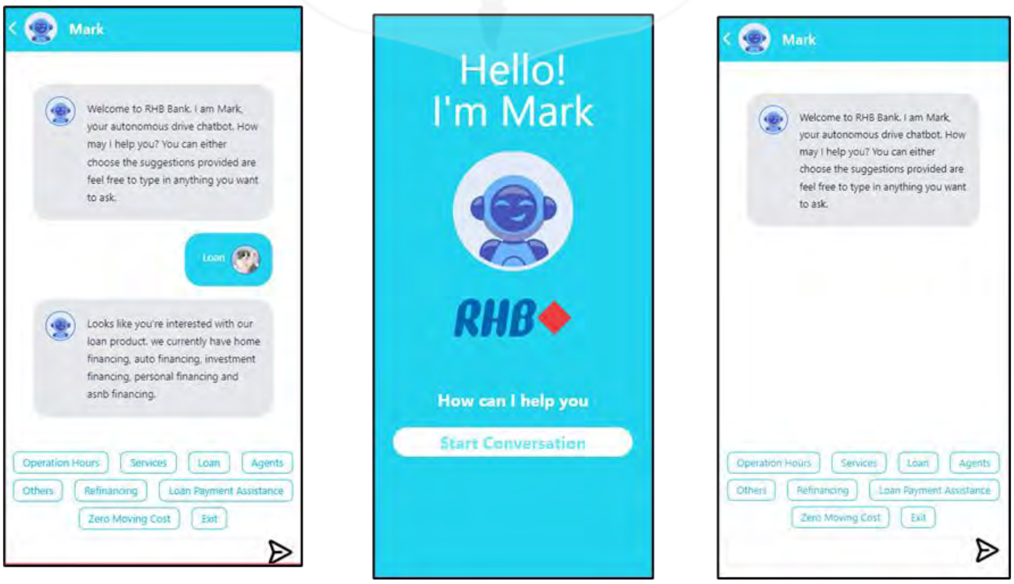


Figure 4.12 Mobile interface of chatbot

4.6 CSAT SCORE CALCULATION

Customer Satisfaction (CSAT) score is a measurement used to quantify the degree to which customers are satisfied. It is implemented in the code to calculate the user's response towards the service provided by a chatbot. The codes were designed to mark the user responses into five (5) categories namely: unsatisfied, very unsatisfied, neutral, satisfied, and very satisfied as shown in Table 4.2. At the end of the conversation with the chatbot, users or customers are prompted to input their satisfaction level by clicking on the button provided. Birkett (2023) defined Customer Satisfaction Score (CSAT) as a medium of measurement towards user's experience when using the provided services. CSAT score will be used to highlight user's sentiment reflected in polarity value as positive, neutral, negate, and compound, with a mixture of all three sentiments. Table 4.2 shows an example of the source code used to indicate the CSAT score.

Table 4.2 Source code used to indicate CSAT Score

CSAT Score Source code

```
processFinalData(user_inputs:list):
    user_response = user_inputs
    sentiment_analysis = calculateSentiment(user_response[:-1])
    csat_score = user_response[-1]
    score = 0
    if(csat_score == "Very Unsatisfied"):
        score = 1
    elif(csat_score == "Unsatisfied"):
        score = 2
    elif(csat_score == "Neutral"):
        score = 3
    elif(csat_score == "Satisfied"):
        score = 4
    elif(csat_score == "Very Satisfied"):
        score = 5
    print("sentiment_analysis: ",sentiment_analysis)
```

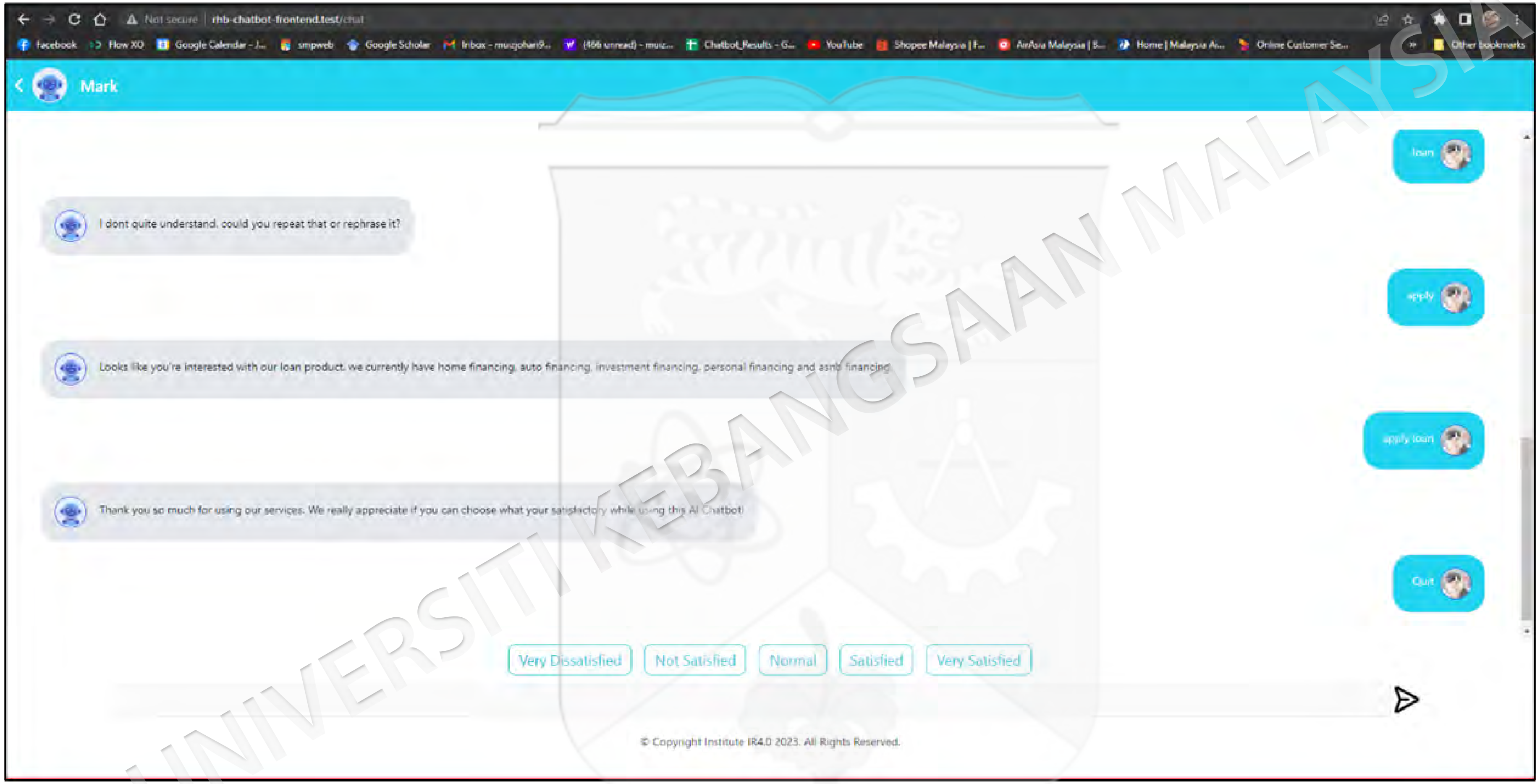


Figure 4.13 CSAT score display in the chatbot

Figure 4.13 illustrates the simple, solid design which is based on the Information Seeking Behaviour Theory and Customer Behaviour theory of users in the banking domain, considering the nature of the banking world.

Figure 4.13 also shows the CSAT score display in the chatbot. The display of the CSAT score is clear and prominent for the customers to see. The options for users or customers to select are from "very dissatisfied" until "very satisfied." Customer transparency and trust are aspects implicitly addressed by the Information Seeking Behaviour (ISB) theory and the Customer Behaviour (CB) theory.

4.7 CONCLUSION

The overall methodology of design and development of the new chatbot (NURUL-C), is based in the Holistic Evolutionary Iterative Humanistic Chatbot Life Cycle (HEI-HC-LC). Based on this methodology and the design and development chatbot framework, details of the activities of each of the three (3) phases are namely: Phase I: Planning Pre-Development (Formative Analysis) EIC-LC; Phase II: Development, Integration, and Implementation EIC-LC; and Phase III: Evaluation and Testing (Summative Analysis) were discussed. Apart from that, the chatbot integration and flow was also discussed based on the front-end and back-end model framework. The context level diagram of the chatbot was also discussed. Important now too, is the discussion on the user interface adopted in the design and development of the chatbot. It was clear that the approach used was to ease use, adopting simple, solid, and clean approach that is in line with the Information Seeking Behaviour (ISB) and Customer Behaviour (CB) theories as well as the outlook and nature of the banking world that portrays simplicity, transparency, trust, and professionalism.

CHAPTER V

FINDINGS AND DISCUSSIONS

5.1 INTRODUCTION

This chapter discusses the findings of the research related to the development of the improved new chatbot (NURUL-C) based on the sentiment analysis technique and the assessment of the Customer Satisfaction Index (CSI). The findings of the study cover five (5) main aspects related to the research objective of the study as follows:

i. Develop the Holistic Evolutionary Iterative Humanistic Chatbot Life Cycle (HEI-HC-LC).
ii. Determine the quality attributes for a good chatbot (in other words to determine quality attributes).
iii. Develop the text analysis technique.
iv. Explore and develop the sentiment analysis for text from the chatbot.
v. Validate users' or customers' level of satisfaction with the chatbot's text analyser.

5.2 FINDINGS ON THE DEVELOPMENT OF THE HOLISTIC EVOLUTIONARY ITERATIVE HUMANISTIC CHATBOT LIFE CYCLE (HEI-HC-LC)

The overall methodology development model is as indicated in Figure 3.1, Chapter III, and is called the Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle (HEI-HC-LC). The process model consists of three (3) main phases. Phase I involves Formative Analysis, preliminary study, or a pilot study to determine the attributes of a good quality chatbot to be applied to the improved new chatbot (NURUL-C). Phase II includes designing the interface based on the System Requirement Specification (SRS), the front-end and back-end framework, the visual elements as well as the energy efficiency elements. This phase also involves integrating the new attributes, developing the text analyser technique, and implementing at the customer's or user's site, includes sentiment analysis for the text of the chatbot by the users' or customers to see if it meets their needs. Phase III involves the Evaluation and testing (Summative Analysis) to assess the level of users' or customers' satisfaction with the interface and the chatbot's text analyser.

5.3 FINDINGS OF THE USER INTERFACE

Formative analysis conducted on the user interface of the chatbot designed and developed. The findings as indicated in Section 4.5, Chapter IV, showed an incredibly positive result. The design was clean, solid, and simple underpinned by the Information Seeking Behaviour (ISB) theory and the Customer Behaviour (CB) theory. It aligned with the straightforward, transparent and trusting characteristics of the banking domain dominated by professionalism. Print screens as shown in the same section, Chapter IV.

5.4 FINDINGS OF OBJECTIVE 1: (RO1): DETERMINE ATTRIBUTES FOR GOOD QUALITY CHATBOT

It found there were two (2) methods used to achieve the first objective namely: literature review and conducting a pilot study. Based on the literature review conducted, there were six (6) prominent attributes commonly included in the design and development of a chatbot. The categories of features or attributes found were functionality, efficiency, technical satisfaction, humanity, effectiveness, and ethics. Table 5.1 shows the summary of the quality attributes acquired from the literature review conducted.

Table 5.1 Summary of quality attributes acquired from literature review

Category	Quality Attributes
Functionality	• Interpret commands accurately (Morrissey and Kirakowski 2013) • Flexible in interpreting knowledge (Cohen & Lane 2016) • Number of services available in the chatbot (Eeuwen 2017) • Suitability (ISO/IEC 2001) • Accuracy (ISO/IEC 2001) • Interoperability (ISO/IEC 2001) • Compliance (ISO/IEC 2001) • Security (ISO/IEC 2001)

to be continued…

…continuation

Efficiency	• Ease of use (Candela 2018 & Duijst 2018) • Quick responds vs free text (Duijst 2018) • Always available (Wang et. al. 2018) • Accessibility (Duijst 2018) • Time behaviour (ISO/IEC 2001) • Resource (ISO/IEC 2001) • Utilization (ISO/IEC 2001) • Strong to manipulation of data input by users (Saaty 1990)
Technical Satisfaction	• Able to convey greeting (Saaty 1990) • Provide emotional information using tones and expression (Saaty 1990)
Humanity	• Realness of robot (Ramos 2017) • Create an enjoyable interaction (Morrissey & Kirakowski 2013) • Convey personalities (Morrissey & Kirakowski 2013; Kuligowska 2015) • Able to maintain the theme of the discussion (Smutny & Schreiberova 2020) • Able to respond to specific inquiry (Smutny & Schreiberova 2020) • Ability to mimic human personality (Saaty 1990) • Ability to provide correct response to user's request (Saaty 1990)
Effectiveness	• Linguistic accuracy (Saaty 1990) • Consists of wide array of knowledge (Saaty 1990)
Ethics	• Trained with knowledge cultural and ethics of users (Saaty 1990) • Sensitivity to user's concern (Saaty 1990)

The determination of the attribute for a decent quality chatbot was based on the instrument developed and administered to samples during the preliminary or pilot study. The instrument, called the AGQC-PS questionnaire validated by six (6) experts in the field of HCI, Chatbot and Banking domain SME. Figure 5.1 shows the Holistic Attribute Analysis Model for Chatbot Demographic Info (HAAM-C-1). Figure 5.2 shows the Demographic Information and their experiences as well as opinion on features of a quality chatbot.

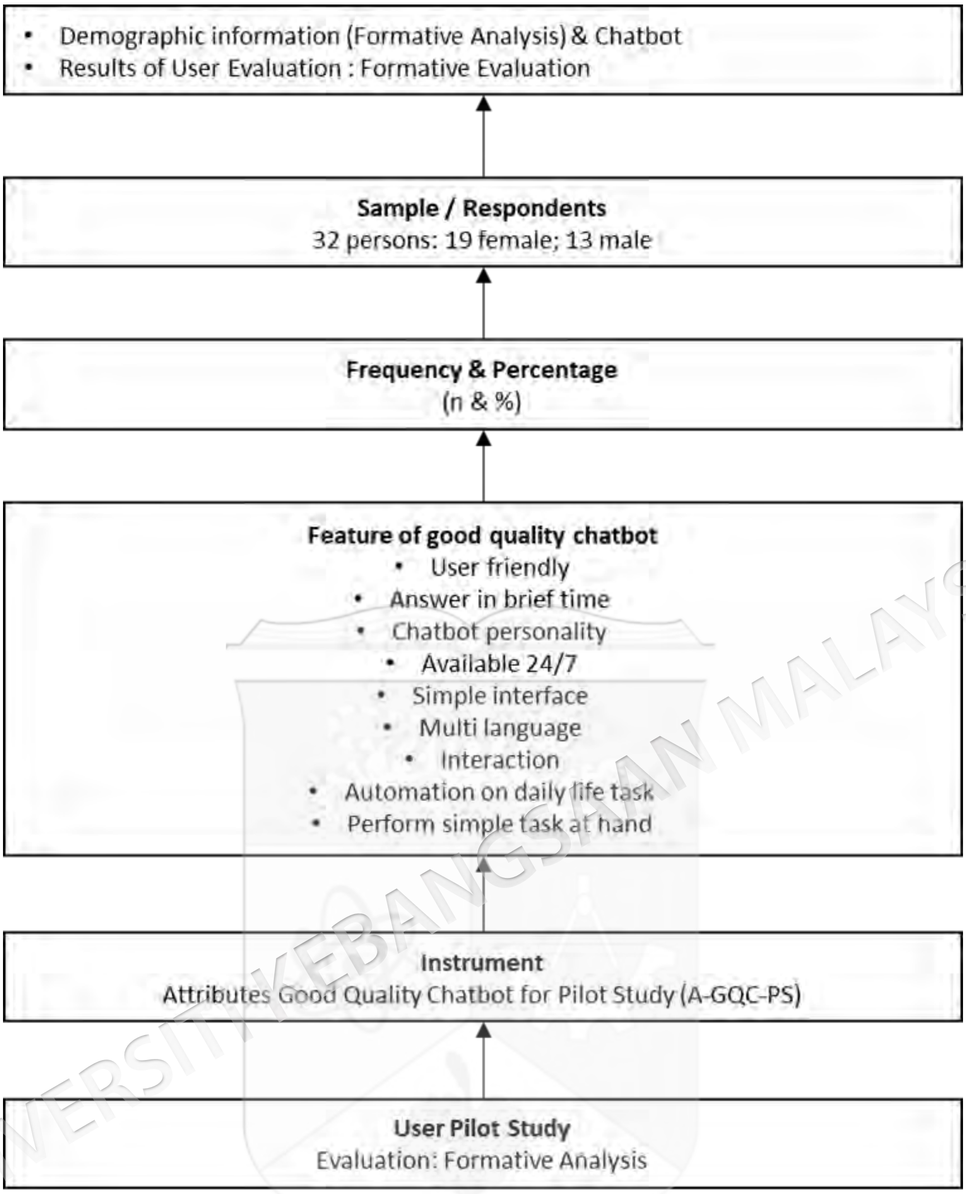


Figure 5.1 Holistic Attribute Analysis Model for Chatbot (HAAM-C-1): Demographic information

Based on the preliminary study or the pilot study conducted, the findings were based on the sentiment of users' or customers towards the chatbot, and its service, through the questionnaire administered to the users who experienced with chatbots. The inquiry done to identify the user's evaluation and satisfaction level with a current available chatbot. There were three (3) sections in the questionnaire which required to answer. The sections were demographic information, chatbot evaluation and user satisfaction level evaluation. Section A of the instrument was the demographic information as in Figure 5.2.

a. Section A: Demographic Information

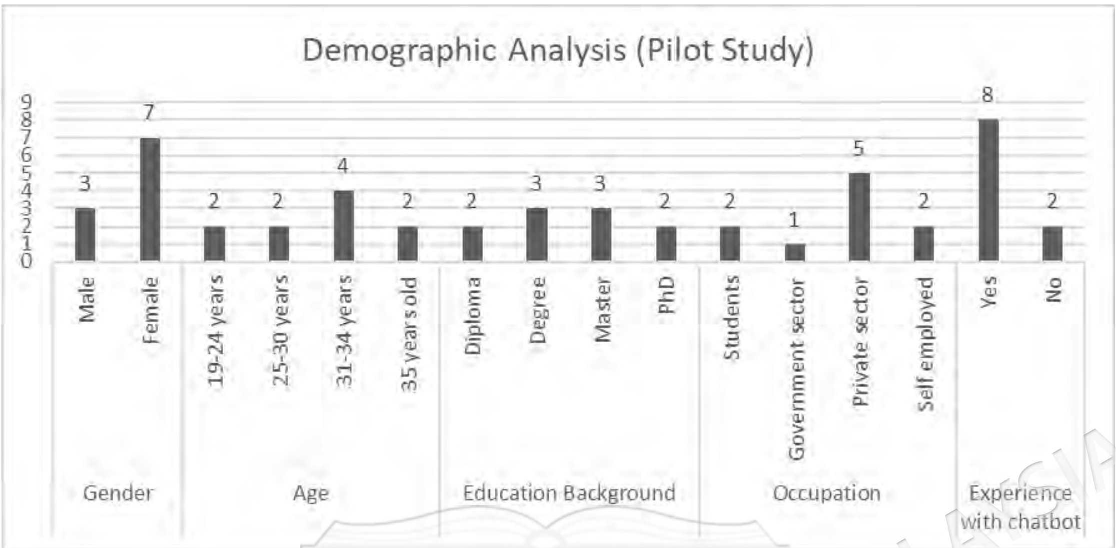


Figure 5.2 Demographic analysis (pilot study)

Figure 5.2 shows that thirty-two (32) samples gave their feedback through Google Forms. Nineteen (19) (59.4%) were female while thirteen (13) (40.6%) of the samples were male. The samples were selected based on purposive sampling as well as randomly despite their gender or other demographic information. Based on the statistics, four (4) of the samples (12.5%) were nineteen (19) to twenty-four (24) years old, fourteen (14) (43.8%) were twenty-five (25) to thirty (30) years old, seven (7) (21.9%) were thirty-one (31) to thirty-four (34) years old, and seven (7) (21.9%) aged thirty-five (35) years old and above. Most of the samples were from the private sector which represented twenty-three (23) (71.9%), followed by four (4) students (12.5%), three (3) government sector (9.4%) and two (2) self-employed (6.3%). The findings from the descriptive analysis showed that twenty-six (26) samples were familiar with Siri, seven (7) were familiar with Cortona, five (5) were familiar with Air Asia AVA, four (4) were familiar with Alexa, CIMB EVA and Celcom – Emma & Clive while one (1) were familiar with Google Mic and Google. Based on the descriptive analysis, there are twenty-six (26) samples who chose “user-friendly” as the most important feature of a chatbot, followed by the “answer in short time” (24 respondents), “available 24/7” (23 respondents), “simple interface” (21 respondents), “interaction” (19 respondents), “multi-language” (18 respondents), “chatbot personality” (14 respondents), and one (1) participant for “automation on daily life task” and “perform simple task at hand”.

b. Section B: Findings of Chatbot Evaluation: Effectiveness of Chatbot

Section B of the instrument (A-GQC-PS) or the Attribute Good Quality Chatbot for the pilot study consists of the chatbot evaluation. This section consists of ten (10) statements which evaluate the effectiveness of Chatbot based on the users' or customers' perspective. Samples were asked to select one of the options available in the Likert Scale provided which range from; (1) Strongly disagree, (2) Disagree, (3) Neutral, (4) Agree, and (5) Strongly agree, based on the Holistic Attribute Analysis Model for Chatbot (HAAM-C-2) on Effectiveness of the chatbot as in Figure 5.3. Figure 5.4 explains the results of analysis on the effectiveness of chatbot in the chatbot evaluation instrument.

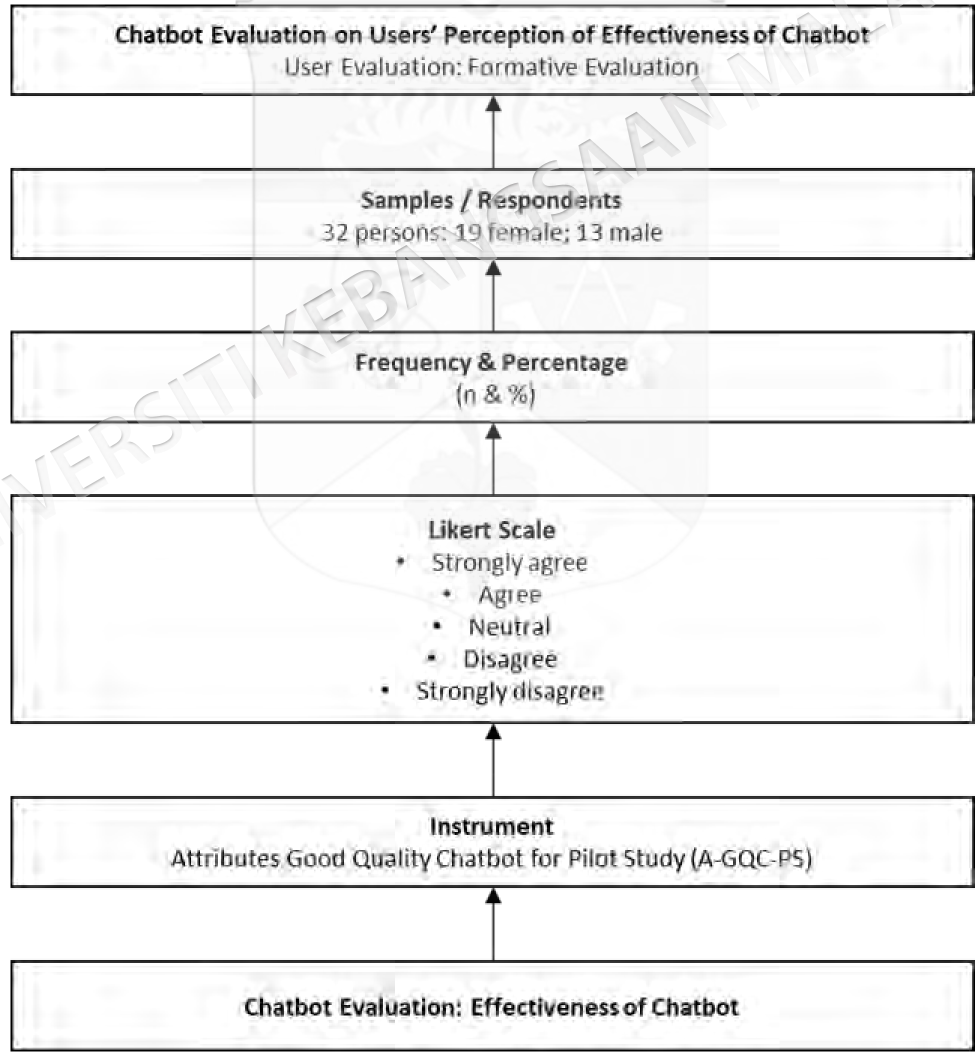
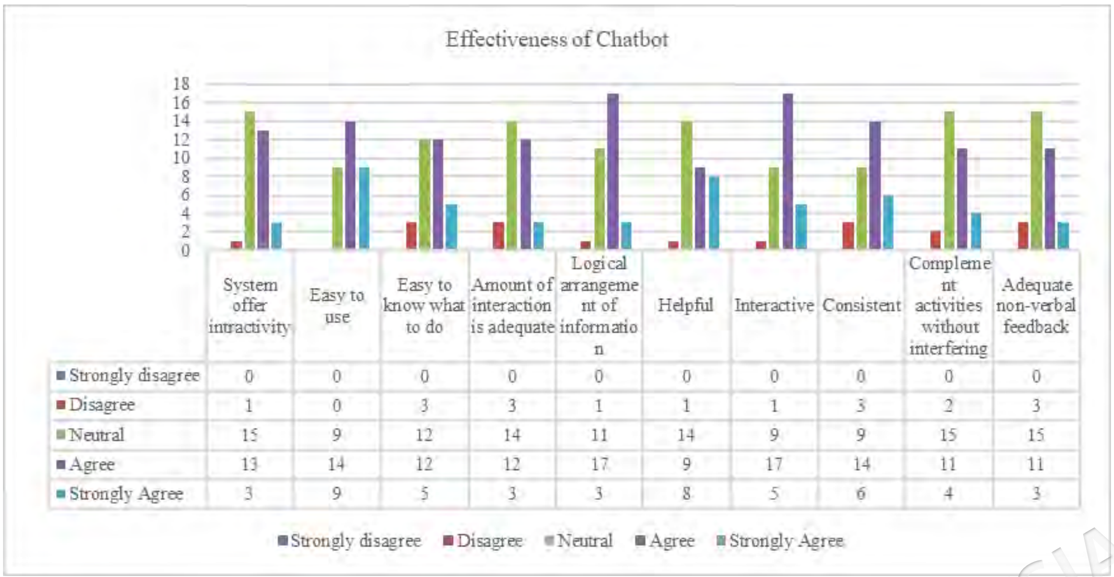


Figure 5.3 Holistic Attribute Analysis Model for Chatbot (HAAM-C-2): Chatbot evaluation on user's perception on effectiveness of chatbot



	System offer intractivity	Easy to use	Easy to know what to do	Amount of interaction is adequate	Logical arrangement of information	Helpful	Interactive	Consistent	Complement activities without interfering	Adequate non-verbal feedback
Strongly disagree	0	0	0	0	0	0	0	0	0	0
Disagree	1	0	3	3	1	1	1	3	2	3
Neutral	15	9	12	14	11	14	9	9	15	15
Agree	13	14	12	12	17	9	17	14	11	11
Strongly Agree	3	9	5	3	3	8	5	6	4	3

Figure 5.4 Effectiveness of chatbot

In general, results of the findings in Figure 5.4 showed that current systems or chatbots available did not fully satisfy the samples who were involved in the study. Out of the ten (10) statements stated in the questionnaire only two (2) statement received a 50% agreement namely on statement 5: ('The arrangement of information on the screen is logical') and statement 7: ('The chatbot is interactive'). Two (2) statements were above 40% agreement namely on statement 2: ('The system is easy to use') and statement 8: ("The chatbot reacts in a consistency way"). The rest of the statement received an agreement score between 26% - 35%. None of the statement received a strongly agree statement of more than 26%.

c. Section C: Findings Customer Satisfaction Level on Chatbot Evaluation

Section C of the instrument (A-GQC-PS) or the Attribute Good Quality Chatbot for the Pilot Study consists of the Customer Satisfaction Level Evaluation. This section consists of ten (10) statements which evaluate the customer's satisfaction level using the Chatbot. Participants were required to choose from the options available in the Likert Scale provided which range from; (1) Strongly disagree, (2) Disagree, (3) Neutral, (4) Agree, and (5) Strongly agree. Figure 5.6 summarises the analysis for this section based on the Holistic Attribute Analysis Model for Chatbot (HAAM-C-3): Customer Satisfaction Level on Chatbot Evaluation as in Figure 5.5.

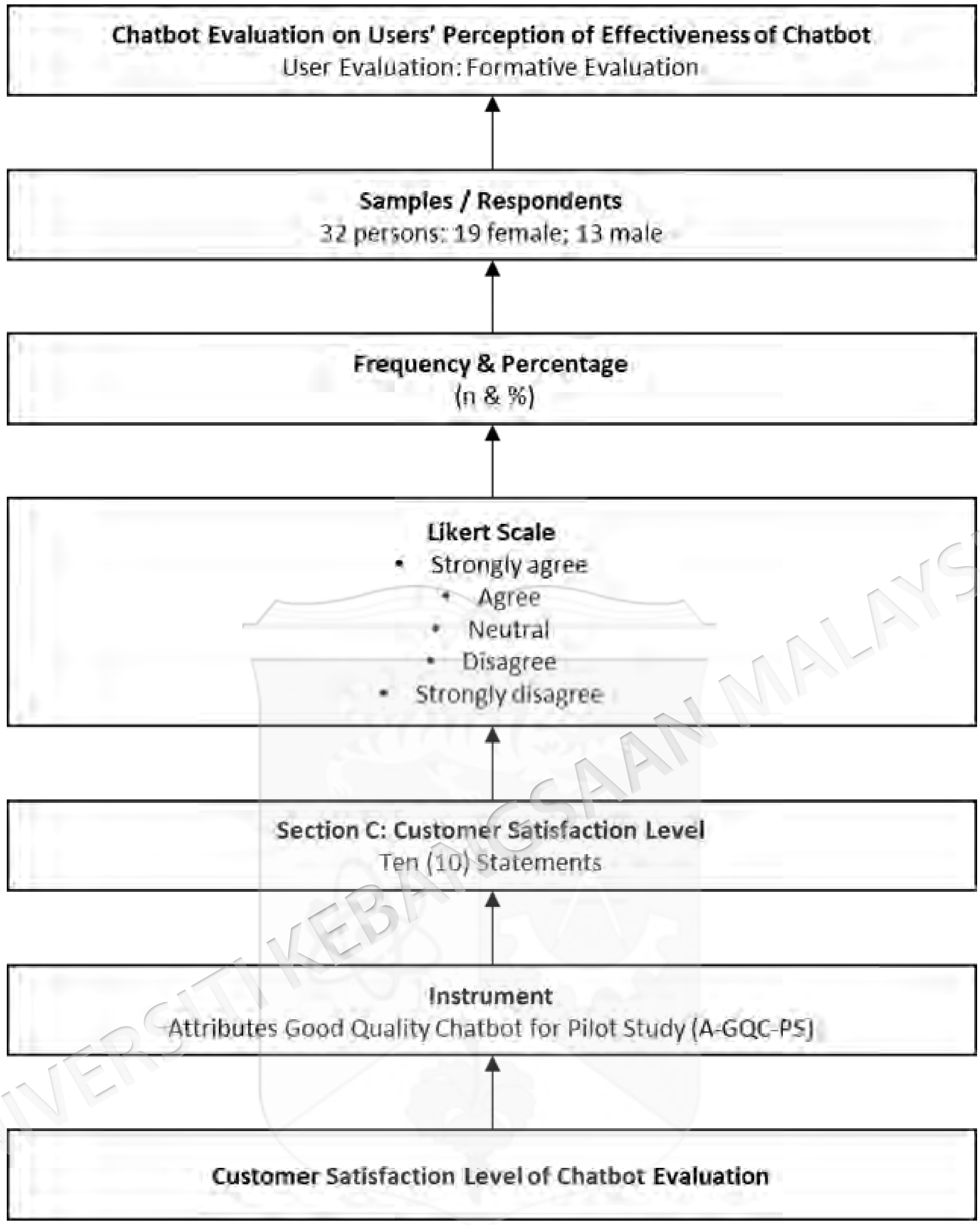
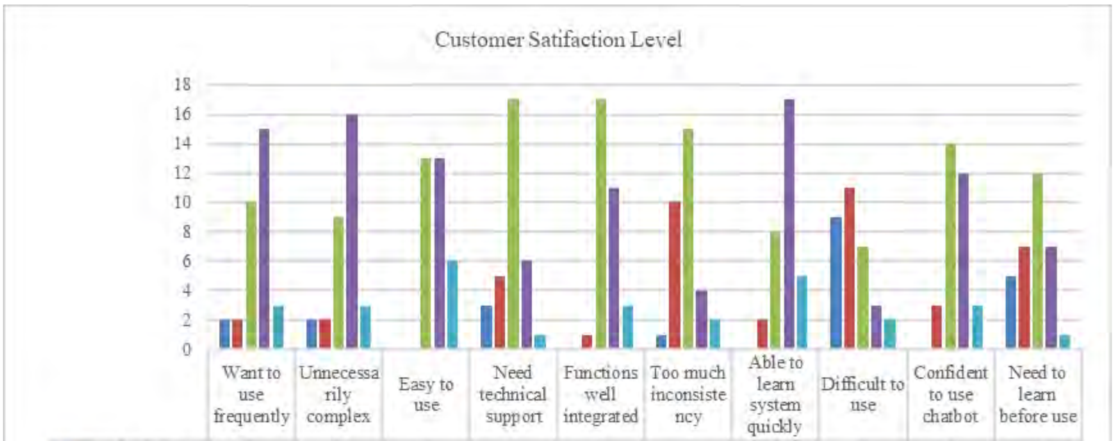


Figure 5.5 Holistic Attribute Analysis Model for Chatbot (HAAM-C-3): Customer satisfaction level on chatbot evaluation



	Want to use frequently	Unnecessarily complex	Easy to use	Need technical support	Functions well integrated	Too much inconsistency	Able to learn system quickly	Difficult to use	Confident to use chatbot	Need to learn before use
Strongly disagree	2	2	0	3	0	1	0	9	0	5
Disagree	2	2	0	5	1	10	2	11	3	7
Neutral	10	9	13	17	17	15	8	7	14	12
Agree	15	16	13	6	11	4	17	3	12	7
Strongly Agree	3	3	6	1	3	2	5	2	3	1

Strongly disagree ■ Disagree ■ Neutral ■ Agree ■ Strongly Agree

Figure 5.6 Customer satisfaction level

In general, results of the findings on the customer satisfaction level in Figure 5.6 on the chatbot evaluation showed that only one (1) out of the ten (10) statements in the questionnaire received a 50% agreement on statement 7: ('I would imagine most people would learn to use this system very quickly'). For most of the statement (>35% - 50%) were either neutral or disagree.

Therefore, based on the evaluation conducted for the Section B and Section C, it can be inferred that the samples who represent the users or customers of the chatbot were not totally satisfied with the current attributes of the chatbots. There was a need to design and develop a new improved chatbot based on some new attributes which will be discussed in detail in Section 5.5 of this thesis.

5.5 FINDINGS OF OBJECTIVE 2 (RO2): FINDINGS ON THE DEVELOPMENT OF TEXT ANALYZER TECHNIQUE

The development of the chatbot assisted by the integration of procedures for data pre-processing. Once the features are determined after answering the first objective, next in place is to decide the text analyzer techniques. The programming language used for the development of the back-end scripts was Python. The flow included tokenisation, stopword removal, and text normalization which are stemming and lemmatisation. The text input by selected participants can be in various forms such as with typing errors, usage of emoticons or icons, multi-language as well as in short forms.

Tokenisation was used in the back end of the chatbot to reduce large sentences into smaller bits known as tokens. This can help to segregate the sentence into groups which is known as word tokenisation (Pandey et al. 2019). The study uses nltk word_tokenizer () to split the sentence into words.

The process then continues with stopwords removal. During this stage, any unimportant word is removed which as “is,” “the,” “a” and others. As shared by Pandey et al. (2019), the scripts for removing stop words and punctuation in text data are as:

```
import math
from nltk.corpus import stopwords
self=[ "`","``",""""]
stop_words = set(stopwords.words(“english”) + list (string.punctuation) + list(self))
```

Next, the text continued to undergo the process of text normalisation. There were two (2) processes happening in this process which were stemming, and lemmatisation as indicated in the Holistic Attribute Analysis Model for Chatbot (HAAM-C-4): Development of Text Analyzer Technique seen in Figure 5.7.

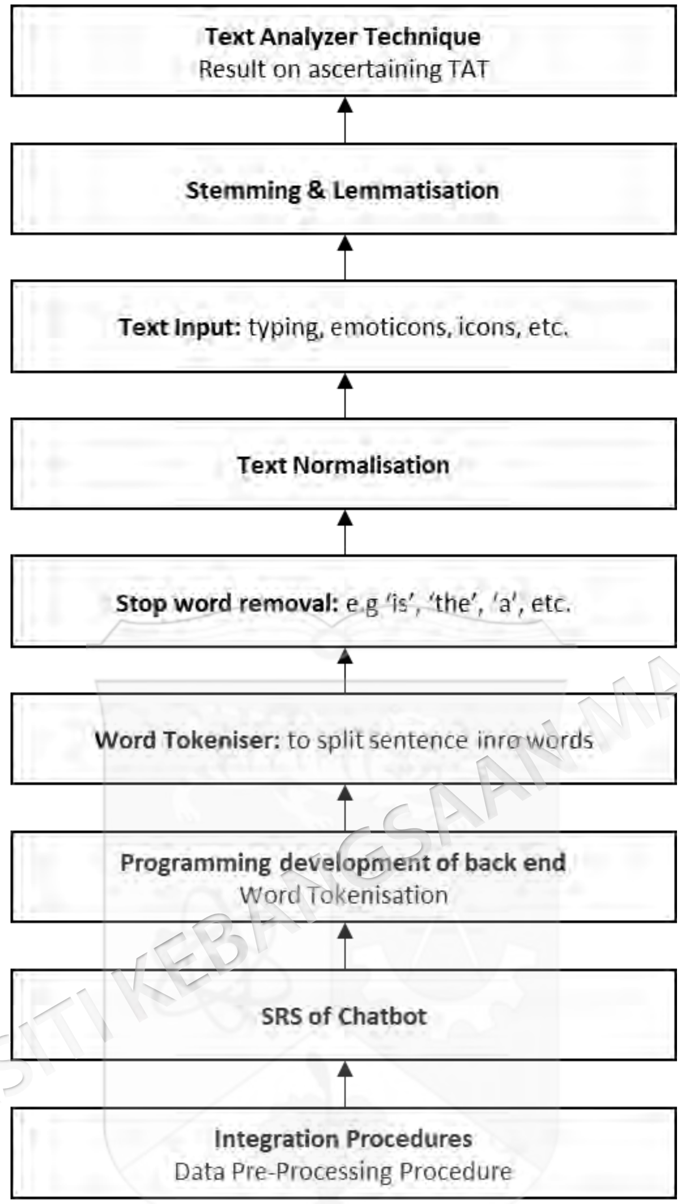


Figure 5.7 Holistic Attribute Analysis Model for Chatbot (HAAM-C-4): Development of text analyser technique

5.6 FINDINGS OF OBJECTIVE 3 (RO3): FINDINGS OF SENTIMENT ANALYSIS FROM TEXT ANALYSIS IN CHATBOT

There are two (2) phases of analysis conducted using the scripts obtained from the conversations between the samples and the Chatbot. Based on the conversations, the polarity of the sentiment used is detected in four (4) categories, positive, neutral, negative, and compound.

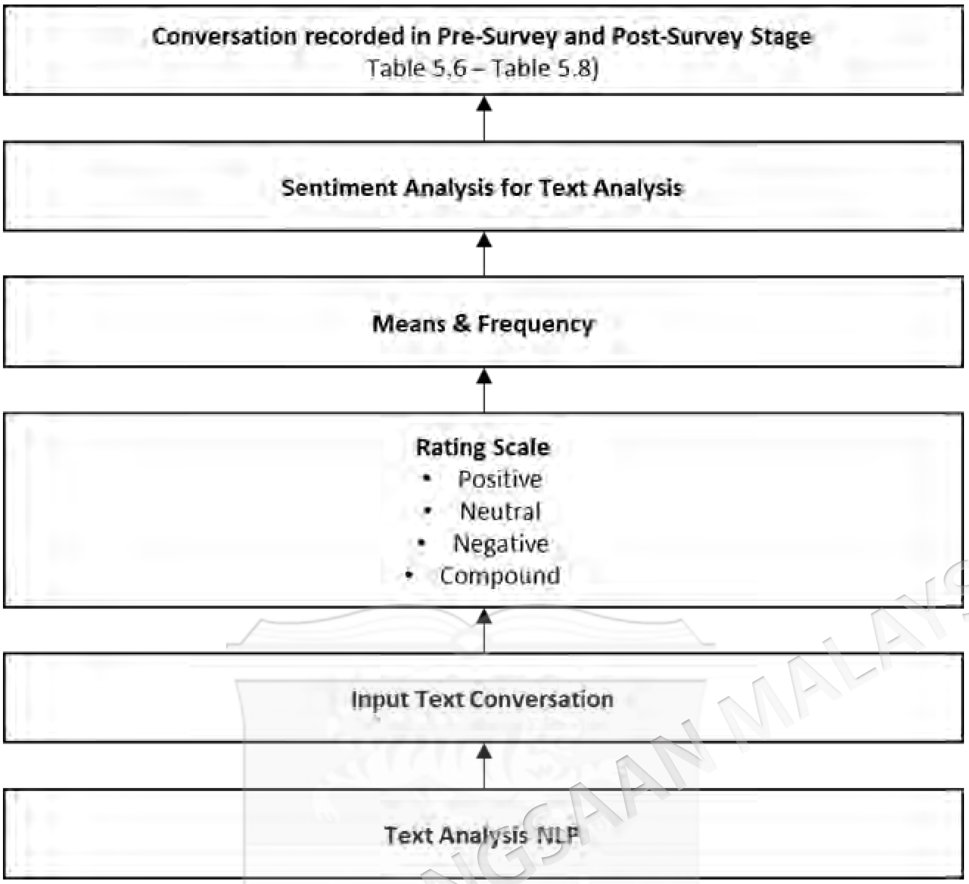


Figure 5.8 Holistic Attribute Analysis Model for Chatbot: Sentiment Analysis (HAAM-C-SA-5)

Figure 5.10 shows the Natural Language Processing (NLP) result for the pre-survey conducted based on the Holistic Attribute Analysis Model for Chatbot (HAAM-C-5), used to conduct the analysis and the model is as indicated in Figure 5.8. Based on the text script, in Table 5.2, the most used word was “loan” which was repeated sixteen (16) times, “service” was repeated seven (7) times, “personal financing” was repeated four (4) times, “card” was repeated four (4) times, “operations” was repeated three (3) times, followed by the words “banking hour”, “live agent”, “loan request’, “nearest branch”, and “variable rate’ which were all repeated twice (2) which visualized as in Figure 5.9. The word cloud created for the conversation at the pre-survey stage is as indicated in Figure 5.9 shows most words used in the conversation recorded in the pre-survey session with six (6) participants.

Figure 5.9 Word cloud for conversation in stage pre-survey

Table 5.2 Conversation I recorded in pre-survey stage

User	Input / Text conversation
1	['Operation', 'Services', 'ghgh', 'where nearest rhb bank', 'Services', 'loans such as', 'perosnal financing such as', 'personal financing such as', 'Quit', 'Very Satisfied']
2	['hi', 'how to start', 'serviecs provided?', 'service provided?', 'types of loan', 'explain on asnb financing', 'Quit', 'Satisfied']
3	['Loan', 'home financing', 'skim rumah pertamaku', 'how many interest u offer for a home loan?', 'Loan', 'auto financing', 'how much is the variable rates?', 'how much is the interest for variable rates?', 'Others', 'can you give me the contact phone number of the loan officer?', 'Quit', 'Neutral']
4	['Personal Financing', 'card credit', 'ceo rhb', 'islamic loan', 'Refinancing', 'Others', 'house loan', 'limit loan', 'Quit', 'Very Satisfied']
5	['I would like to know my balance', 'Where is the nearest branch?', 'Loan', 'Home', 'Operation', 'Others', 'Loan', 'Loan', 'auto', 'Get started', 'Get Started', 'Live Agents', 'Loan', 'Personal Financing', 'Loan', 'ansb', 'asnb', 'Services', 'cards', 'banking hours', 'Services', 'Others', 'hi', 'help', ':(', 'agent', 'cost', 'Loan Request', 'lost of crd', 'lost card', 'Website down time', 'Are you reliable?', "It's very difficult to pay back my loans. Is there a way to reduce payback?", 'Refund', 'security', 'accessibility', 'availability', 'Quit', 'Neutral']
6	['I would like to know my balance', 'Where is the nearest branch?', 'Loan', 'Home', 'Operation', 'Others', 'Loan', 'Loan', 'auto', 'Get started', 'Get Started', 'Live Agents', 'Loan', 'Personal Financing', 'Loan', 'ansb', 'asnb', 'Services', 'cards', 'banking hours', 'Services', 'Others', 'hi', 'help', ':(', 'agent', 'cost', 'Loan Request', 'lost of crd', 'lost card', 'Website down time', 'Are you reliable?', "It's very difficult to pay back my loans. Is there a way to reduce payback?", 'Refund', 'security', 'accessibility', 'availability', 'Quit', 'Neutral']
7	N/A
8	N/A
9	N/A
10	N/A

**N/A: data not available due to technical difficulties faced during conversation recording*

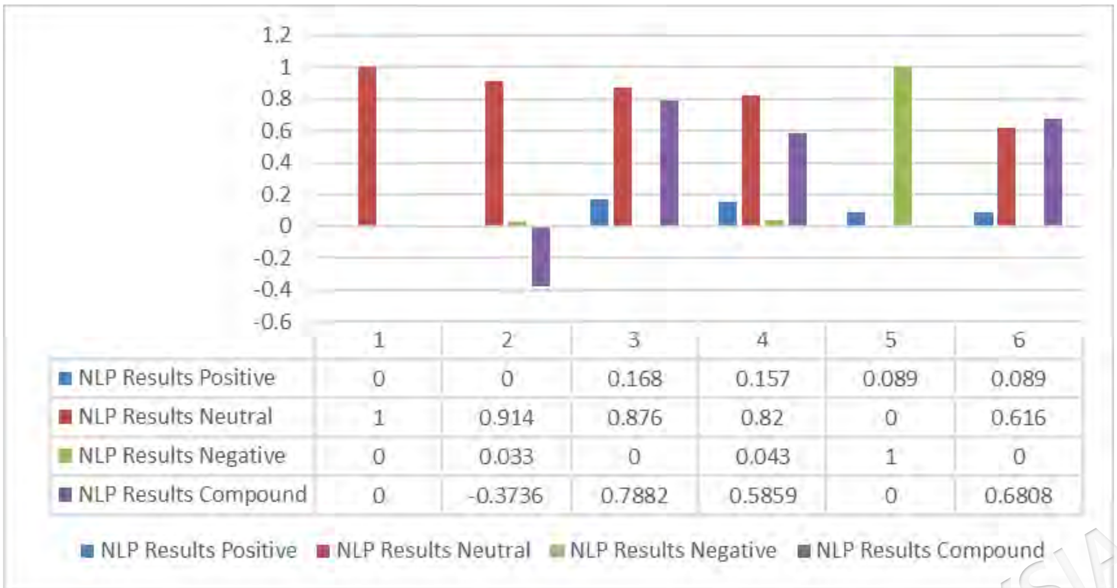


Figure 5.10 Pre-survey NLP result

Figure 5.12 shows the analysis from post-survey NLP result. The result recorded NLP results for ten (10) participants shows usage of neutral and mixed (between positive and negative) in their keyword input into the chatbot.

The word cloud for conversations the post survey stage is as in Figure 5.11. The second stage or the post-survey stage recorded the word “loan” repeatedly by twenty-two (22) times, “service” eight (8) times, “operation” seven (7) times, “zero moving cost” and “refinancing” were repeated five (5) times, both “live agent’ and “mark” were repeated four (4) times, followed by “document” that was repeated for three (3) times, and lastly “tenure for housing loan” and “loan payment assistance” were both repeated twice (2) can be observed in Table 5.3.

Figure 5.11 Word cloud for conversation in post-survey

Table 5.3 Conversation II recorded in post-survey stage

User	Input / Text conversation
1	N/A
2	['Loan', 'what do you suggest?', 'suggestion for personal loan', 'how about loan formy company?', 'can i have loan for my company?', 'private company loan', 'i want to make an investmetn', 'i want to make an investment', 'Zero Moving Cost', 'Loan Payment Assistance', 'Refinancing', 'Refinancing', 'Others', 'Live Agents', 'Services', 'Operation', 'i live in jitra, where is the nearest branch?', 'my bank card is missing, so what should i do?', 'Quit', 'Neutral']
3	loan', 'Loan', 'ASNB', 'Loan', 'Services', 'housing loan', 'current OPR ?', 'OPR', 'current interest rate ?', 'yes', 'how long tenure for housing loan ?', 'what document needed to apply loan ?', 'what meaning of RHB ?', 'what is RHB ?', 'RHest rate ? what is fix rate ? what is floating rate ? how i want to apply loan ? yes how long tenure for housing loan ? what document needed to apply loan ? what meaning of RHB ? what is RHB ? RHB RHB HQ? location of RHB ATM RHB contact ?> Selamat pagi Quit
4	hello', 'what is business loan', 'what are requirement to make a loan', 'can you explain more?', 'what are service that you can provide to me', 'what is differenece debit and credit card', 'oh no', 'Services', 'Loan', 'Refinancing', 'digital banking', 'okay', 'thank you for your servce', 'nothing', 'Quit', 'Satisfied'
5	Operation', 'Services', 'Zero Moving Cost', 'loan', 'Loan', 'Others', 'time', 'product', 'Quit', 'Neutral'
6	'evening', 'Operation', 'security', 'saving', 'save money', 'loan', 'apply', 'apply loan', 'Quit', 'Satisfied'
7	Operation', 'loan interest rate', 'loan settlement', 'asnb financing', 'auto financing', 'auto financing for import car', 'how to check loan balance', 'Live Agents', 'hq address', 'please state neares branch address', 'Zero Moving Cost', 'Services', 'stock broking', 'broking', 'Quit', 'Very Satisfied'
8	Services', 'Loan', 'how i can do loan?', 'auto fanicing loan, how ii can apply?', 'Operation', 'lunch hour operation hours?', 'when operation closed?', 'Live Agents', 'Zero Moving Cost', 'Quit', 'Satisfied'
9	hello mark', 'i want to know about service bank. can you help me mark?', 'i want to learn about premier banking', 'can you tell me about premier banking, mark?', 'what is loan saving', 'what is asnb financing', 'Operation', 'Services', 'Loan', 'Live Agents', 'Others', 'Refinancing', 'Loan Payment Assistance', 'Zero Moving Cost', 'hi mark, can you tell me about loan', 'can i ask something mark?', 'are you an AI ?', 'how about chatgpt', 'can i apply credit card', 'quit', 'Very Satisfied'
10	how are u mark', 'eeeeekkk', 'how are u mark', 'hahaha', 'Others', 'Refinancing', 'personal loan?', 'interest?', 'opr?', 'are you sure not affected by opr?', 'aish mark', 'minimum rate sallary?', 'how about car loan?', 'how to apply for loan?', 'what document i should prepare for loan?', 'who should i contact for the loan details?', 'any specific person?', 'who can apply for load?', 'there must be a guarantor?', 'maximum salary to apply for a loan?', 'branch near me?', 'any related document need to carry?', 'ok mark gud bye', 'Quit', 'Neutral'

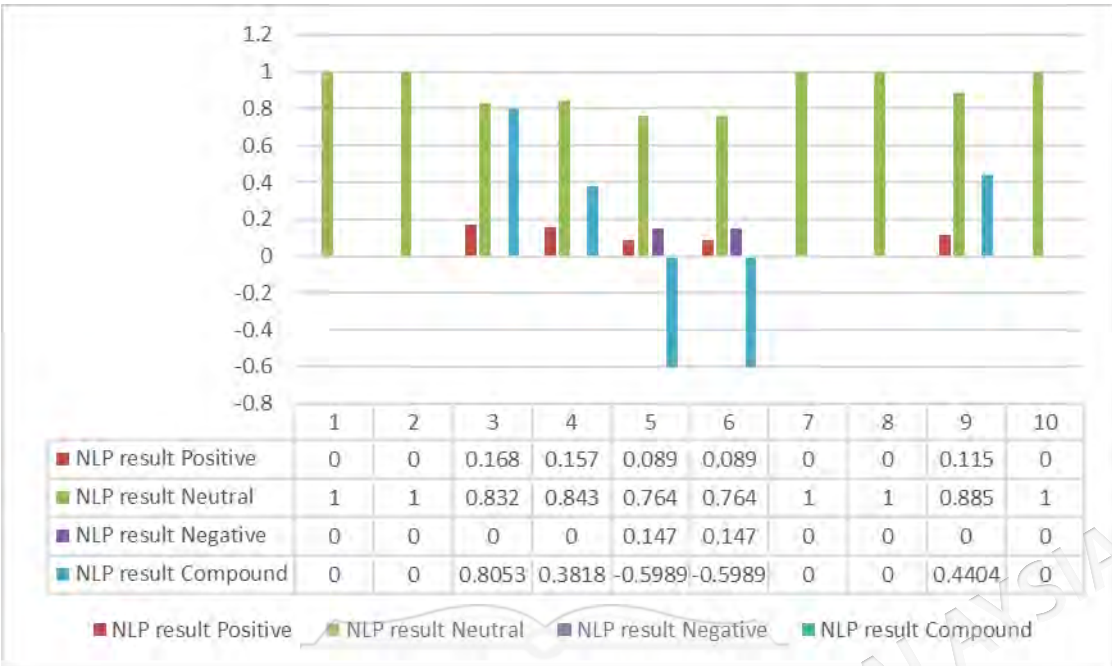


Figure 5.12 Post-survey NLP result

5.7 FINDINGS OF OBJECTIVE 4 (RO4): FINDINGS OF USER'S SATISFACTION LEVEL ON NEW CHATBOT

The level of user's satisfaction was conducted in two (2) phases: pre-survey and post-survey. Both surveys were participated by ten (10) samples who were selected randomly regardless of their background of using chatbots in their daily lives. This section will discuss the outcomes from both surveys which cover their demographic information, and testing scenarios used in the pre-survey phase as well as post-survey. At the end of this section, the researcher will discuss the CSAT score recorded from each session and discuss the effect towards finding the level of performance as well as the relationship between the score and the user's satisfaction level.

5.7.1 Findings in Pre-Survey Analysis

A pre-survey was conducted to gain the first impression of the end user on the new developed chatbot. Ten (10) samples were selected randomly for their opinion on the chatbot. The outcome of this survey was used for the improvement of the chatbot which was later assessed with a larger group of samples. A descriptive analysis was applied using the software named SPSS. The analysis was done based on the Holistic Attribute Analysis Model for New Chatbot (HAAM-NC). Table 5.8 shows the Demographic

Analysis for Pre-Survey. The analysis was done based on the Holistic Attribute Analysis Model of New Chatbot: Demographic Information (HAAM-NC-1) as indicated in Figure 5.13.

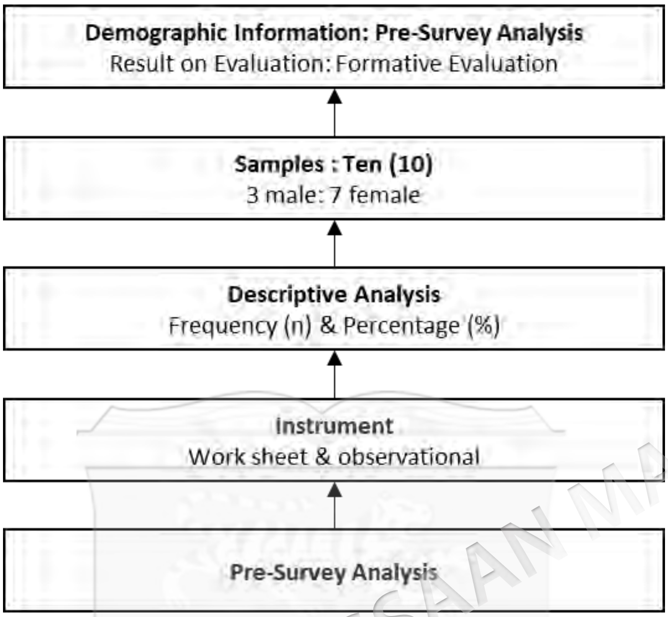


Figure 5.13 Holistic Attribute Analysis Model for New Chatbot: Demographic Information (HAAM-NC-1)

a. Findings on descriptive analysis for demographic information

Based on the analysis, there were three (3) male (30.0%) and seven (7) female (70.0%) samples that contributed to the pre-survey phase. From the total samples, four (4) of them (40.0%) were between the ages of thirty-one (31) and thirty-four (34) years old, and the remaining age group represented two (2) participants comprised 20.0% of the total group. For education background, two (2) participants (20.0%) graduated with diploma, three (3) participants (30.0%) with undergraduate degree, three (3) participants (3%) with master's degree, and two (2) participants (20.0%) with PhD qualification. There were two (2) samples (20.0%) who were students, one (1) (10.0%) was a government sector employee, five (5) samples (50.0%) were private sector employees, and the remaining two (2) samples (20%) were self-employed. As for the experience with using a chatbot, most samples which were eight (8) people (90.0%) were familiar with the application while the remaining two (2) samples (20.0%) were not familiar with using a chatbot. Figure 5.14 summarized the result for demographic analysis.

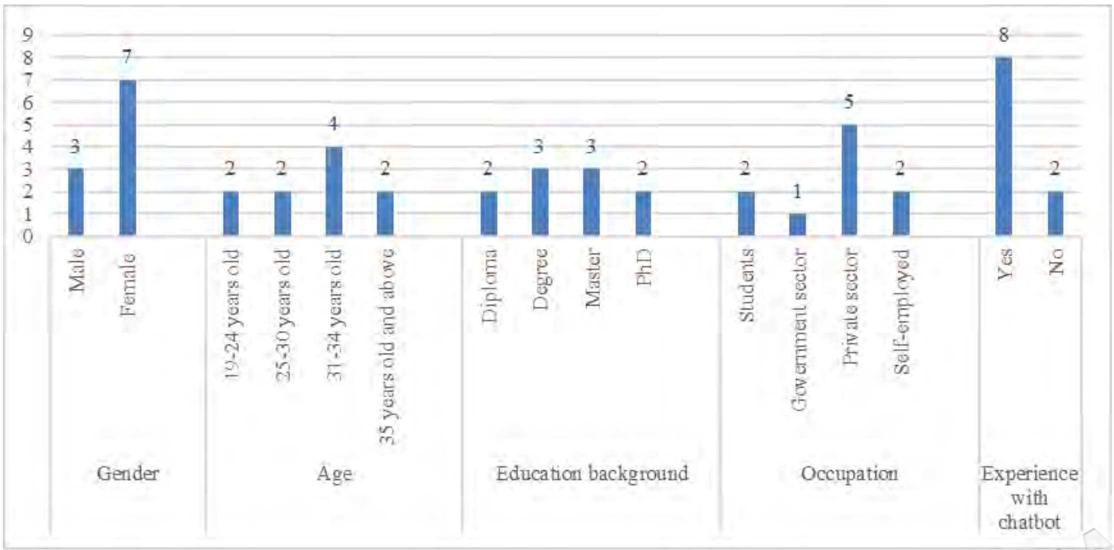


Figure 5.14 Demographic analysis for pre-survey

b. Findings on testing scenario

During the pre-survey activities, the samples were given a sheet of instructions on how to operate the chatbot. Three (3) scenarios were given to each sample to assess and use freely the chatbot and then to leave their comments for further improvement during the post-survey phase. The scenarios given were as follows:

i. Start using the chatbot.

1. Click “Get Started.”
2. Key in any keyword to begin the conversation.
3. Users may use the button for shortcuts.
4. End of conversation.

ii. Find the type of service provided.

1. Click “Get Started.”
2. Key in any related to banking services.
3. Try to input any unrelated keywords.
4. Observe the response.

iii. Find information related to the loan application.

1. Click "Get Started."
2. Key in any keyword related to loan services.
3. End conversation.

In this section, participants were asked to provide their feedback by selecting from the scale provided (1-5; very dissatisfied until very satisfied). There were two (2) items measured as feedback which were "time taken for the chatbot to respond" and the "accuracy of response provided".

The analysis for the following section on usability testing scenario conducted using the Holistic Attribute Analysis Model for New Chatbot: Scenario 1-3 (HAAM-NC-S1-S3) as indicated in Figure 5.15.

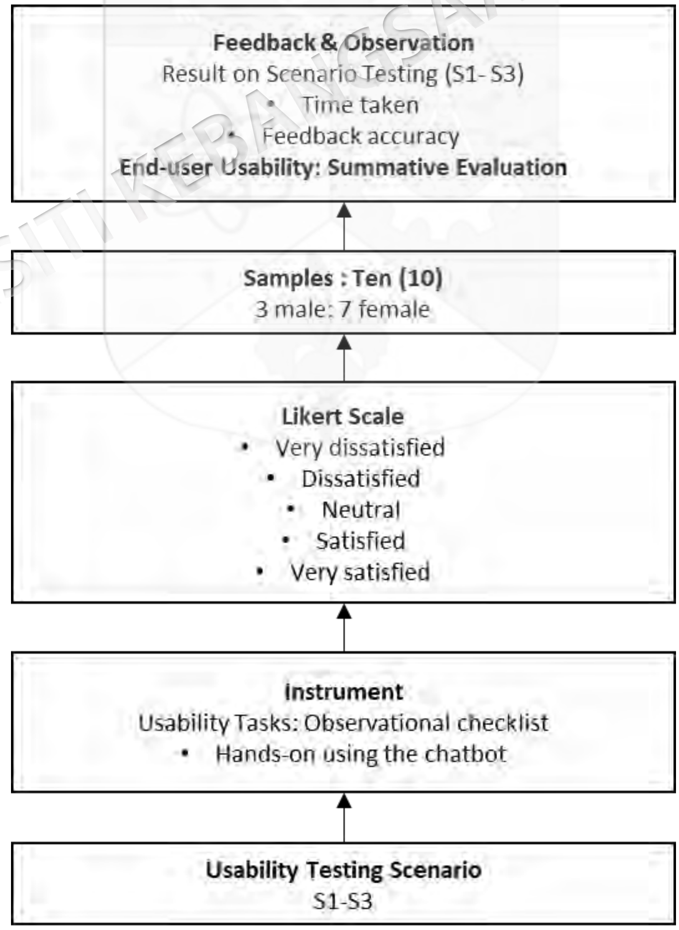


Figure 5.15 Holistic Attribute Analysis Model of New Chatbot: Scenario 1-3 (HAAM-NC-S1-S3)

Table 5.4 shows results on findings of the Scenario-1 Pre-Survey conducted on the new chatbot (NURUL-C). Based on the analysis, 70% of the samples had graded the chatbot with “satisfied” and “very satisfied” and observed it as “pass”. The 30% of samples that were either neutral or dissatisfied commented that the chatbot had “limited available keyword options”; “It takes some time to start a chat after clicking on the ‘Start conversation’”.

Table 5.4 details the feedback received during the pre-survey session for scenario 1. Participants were asked to start using the new chatbot (NURUL-C) unsupervised. They needed to begin the conversation by clicking the provided button or input any keyword to start the conversation. There are seven participants marks “pass” for this activity and three marks “failed.” Feedback was received for the “time taken for the chatbot to respond” (Feedback 1), majority of the participants felt satisfied (three participants) and very satisfied (three participants) with this statement. Two participants felt neutral and the remaining one participant each felt dissatisfied and very dissatisfied. As for the “accuracy of the response provided” (Feedback 2), most of the participants felt neutral about this statement (five participants), three felt dissatisfied, and two participants were satisfied with the accuracy of feedback received from the chatbot. At the end of the session, the participant asked to provide any “remarks for further improvement in the next phase”. Participant Six notes that the limited of keywords available needs to be considered and suggests adding random keywords that relate to services provided by the chatbot. Participant Seven mentioned that “it takes extra time to start the conversation” after clicking the “Start Conversation” button. After checking, the issue was due to internet connection interference which was then resolved after refreshing the application.

Table 5.4 Conversation II recorded in post-survey stage

Participant	**Observation (Pass/Fail)**	**Feedback 1: Time taken for chatbot to respond**	**Feedback 2: Accuracy of response provided**	**Remarks**
1	Pass	Satisfied	Satisfied	
2	Pass	Very Satisfied	Neutral	

to be continued…

…continuation

3	Pass	Very Satisfied	Neutral	
4	Pass	Very Satisfied	Satisfied	Nice work! Using cues.
5	Fail	Dissatisfied	Dissatisfied	
6	Pass	Satisfied	Dissatisfied	Limited to the available keyword options. Still have room for improvement for random keywords (related)
7	Pass	Neutral	Neutral	It takes some time to start chat after clicking on the "start conversation" button
8	Pass	Very Dissatisfied	Neutral	
9	Fail	Satisfied	Dissatisfied	
10	Fail	Neutral	Neutral	

Table 5.5 shows results on the findings of the Scenario-2 Pre-Survey conducted on the new chatbot (NURUL-C). Based on the analysis, 50% of the samples indicated through their observation that the chatbot "passed". The other 50% of the sample were either neutral or very dissatisfied. Among their comment were "It gives too much information about home financing;" "cannot explain about card credit service;" and "system is unable to understand short form."

Table 5.5 Scenario 2 Pre-survey

Participant	Observation (Pass/Fail)	Feedback 1: Time taken for chatbot to respond	Feedback 2: Accuracy of response provided	Remarks
1	Pass	Very Satisfied	Very Satisfied	
2	Pass	Satisfied	Neutral	Does not understand typo
3	Pass	Very Satisfied	Satisfied	It gives the information about home financing and gives the link for further details
4	Pass	Satisfied	Neutral	Cannot explain about card credit service
5	Fail	Satisfied	Very Dissatisfied	1. System could not read a few keywords. For example, 'cards' 2. Add AI into the system
6	Fail	Satisfied	N/A	N/A
7	Fail	Dissatisfied	Dissatisfied	The chatbot is not familiar with the keyword related to banking system

to be continued…

…continuation

8	Fail	Dissatisfied	Dissatisfied	
9	Fail	Satisfied	Dissatisfied	1. Hyperlink not working. 2. Simple keyword, no respond received
10	Pass	Dissatisfied	Neutral	1. Font too small & wording too long. Can reduce length by simplified welcome screen. 2. Proper sentence character 3. Suggest providing button as respond, instead of long sentence / paragraph. 4. Home financing: - Users make typos & system able to response correctly. - Suggest putting action button to click whenever respond not related to user's expectation (i.e., button to speak to live agent or guess close meaning word) 5. When using short form, system unable to understand 6. Hyperlink to website not working.

Table 5.6 shows result on the findings of the Scenario-3 Pre-Survey conducted on the new chatbot (NURUL-C). Based on the analysis, 70% of the samples agreed that through their observation, the chatbot "pass" the test. The 30% of the other samples that were either "dissatisfied" or "neutral" about the chatbot had made comments such as: "did not get answer on interest charged by the bank"; "no basic information on the products offered"; "Failed to detect certain words and constant redirecting to website".

Table 5.6 Scenario 3 Pre-survey

Participant	Observation (Pass/Fail)	Feedback 1: Time taken for chatbot to respond	Feedback 2: Accuracy of response provided	Remarks
1	Pass	Very Satisfied	Very Satisfied	
2	Pass	Satisfied	Satisfied	
3	Fail	Neutral	Neutral	I asked for how many interests did the bank offer but no answer

to be continued…

…continuation

4	Pass	Satisfied	Satisfied	Advice to add more laypeople keyword for public user.
5	Pass	Satisfied	Dissatisfied	Avoid redirecting to website. Provide basic information about the products when asked.
6	Pass	Satisfied	Satisfied	For basic knowledge, the chatbot to works well, but users may opt to go to the main website for information instead.
7	Pass	Satisfied	Satisfied	1. Need to ensure whether this chatbot can function both web and mobile. 2. Font type (size and color) 3. Background color 4. UI (screen, responsive design, chat)
8	Fail	Dissatisfied	Very Dissatisfied	1. Font (type, size, color) 2. Button (color, avoid white font) 3. response take too long. 4. can add a link to the website for further details. 5. Allow the user to provide justification after they select their satisfaction level. 6. Add button “recommend to others”
9	Pass	Satisfied	Neutral	1. Failed to detect words: credit, statement, bank statement. 2. Need to differentiate exit button to avoid confuse
10	Fail	Neutral	Dissatisfied	N/A

c. Findings on New Chatbot (NURUL-C) evaluation

The analysis on the new chatbot evaluation during the pre-survey done based on the Holistic Attribute Analysis Model for New Chatbot (HAAM-NC-E) is as indicated in Figure 5.16.

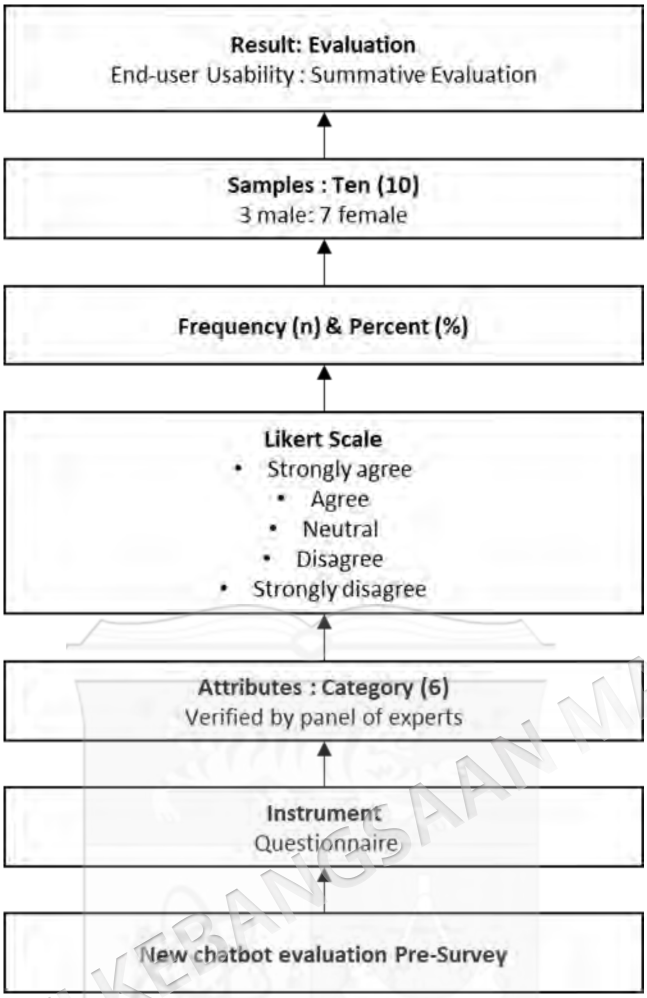


Figure 5.16 Holistic Attribute Analysis Model for New Chatbot Evaluation (HAAM-NC-E)

During pre-survey, samples requested to evaluate the new chatbot, NURUL-C, by answering the questionnaire administered. The measurement was based on the Likert Scale. Figure 5.17 shows the results of the findings obtained from the feedback of the sample.

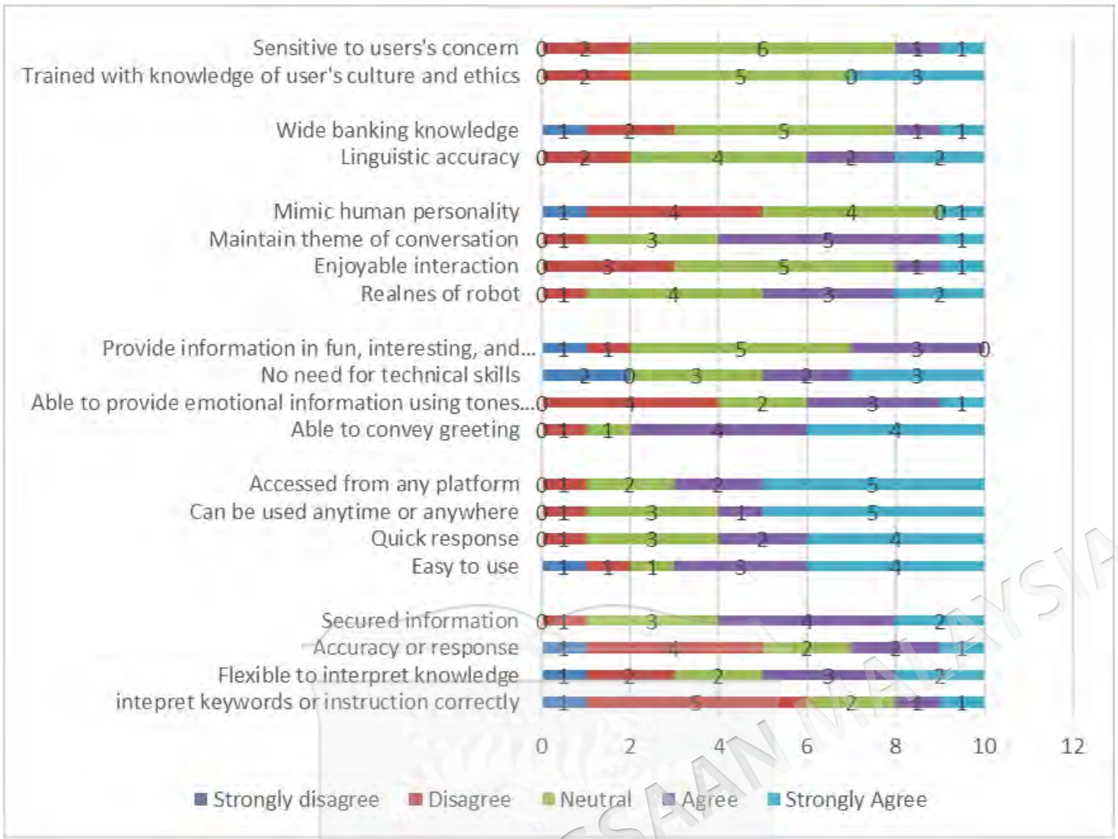


Figure 5.17 New Chatbot evaluation (pre-survey)

The evaluation identified six (6) categories of quality attributes, verified by a panel of experts, with corresponding statements that the survey participants had to answer using the provided Likert Scale. Generally, it can be seen that for the first category on “Functionality”, more than 50% of the samples disagreed or were neutral; for the category on “Efficiency” more than 70% agreed that the chatbot was efficient; on the “Technical Adequacy” category, more than 40% agreed that the chatbot were either neutral or in disagreement; for the fourth category on “Humanity”, more than 50% agreed and more than 10% were is disagreement; for the “Effectiveness” category, more than 30% were agreeable that the chatbot was effective; and lastly, on the “Ethics” category, more than 60% were neutral on the matter.

d. Findings on Customer Satisfaction Index (CSI)

The analysis on the Customer Satisfaction Index (CSI) was based on the Holistic Attribute Analysis Model: CSI (HAAM-NC-CSI) as indicated in Figure 5.18. The last section asked in the questionnaire was related to the Customer Satisfaction Index (CSI) evaluation. Figure 5.19 recorded all the details and feedback received from the participants.

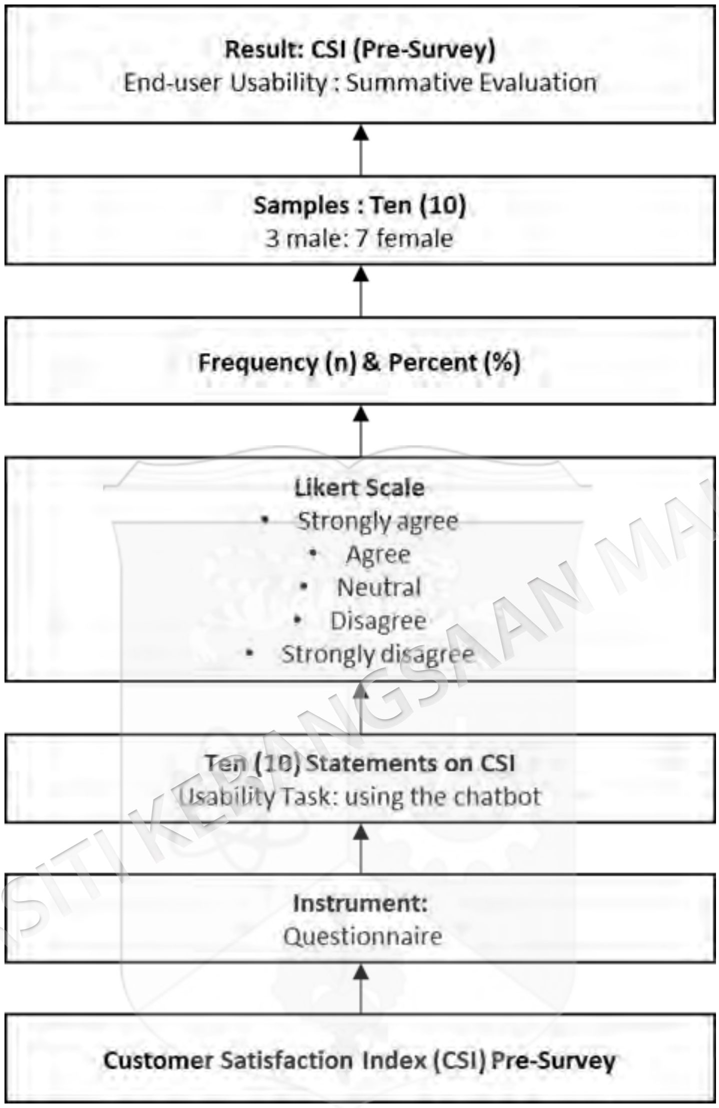


Figure 5.18 Holistic Attribute Analysis Model: Customer Satisfaction Index (HAAM-NC-CSI)

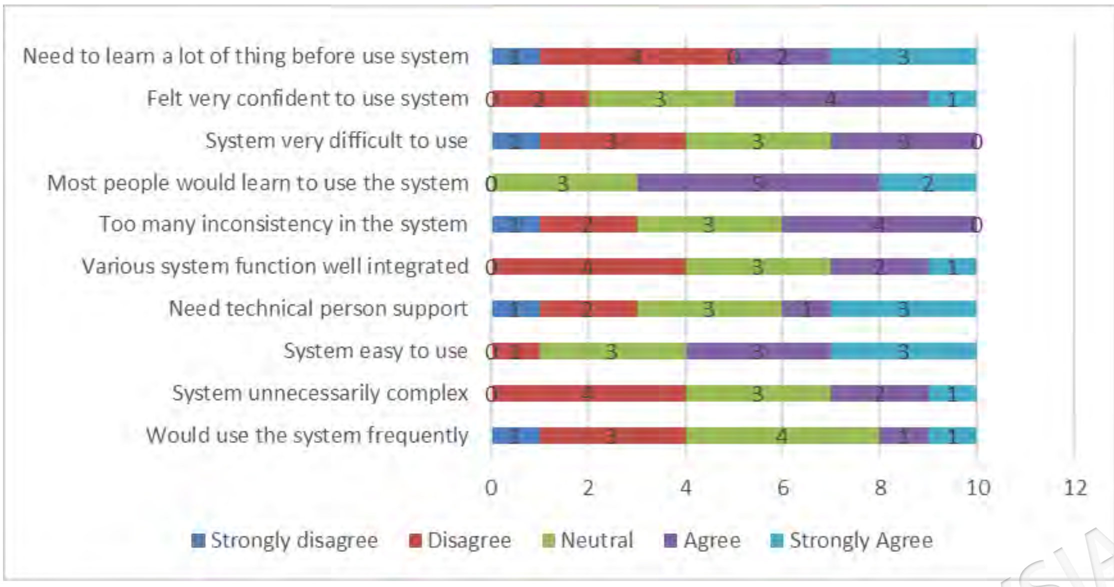


Figure 5.19 Customer Satisfaction Index (CSI) pre-survey

5.8 FINDINGS ON COMPARISON OF USER'S LEVEL OF SATISFACTION TOWARDS CHATBOT AND IMPROVED CHATBOT

Findings of the comparison of users' level of satisfaction towards the chatbot and the improved or new chatbot, NURUL-C, was done based on the Holistic Attribute Analysis Model: Users' Level of Satisfaction between Chatbot and New Improved Chatbot (HAAM-ULS-C&NC) as indicated in Figure 5.20.

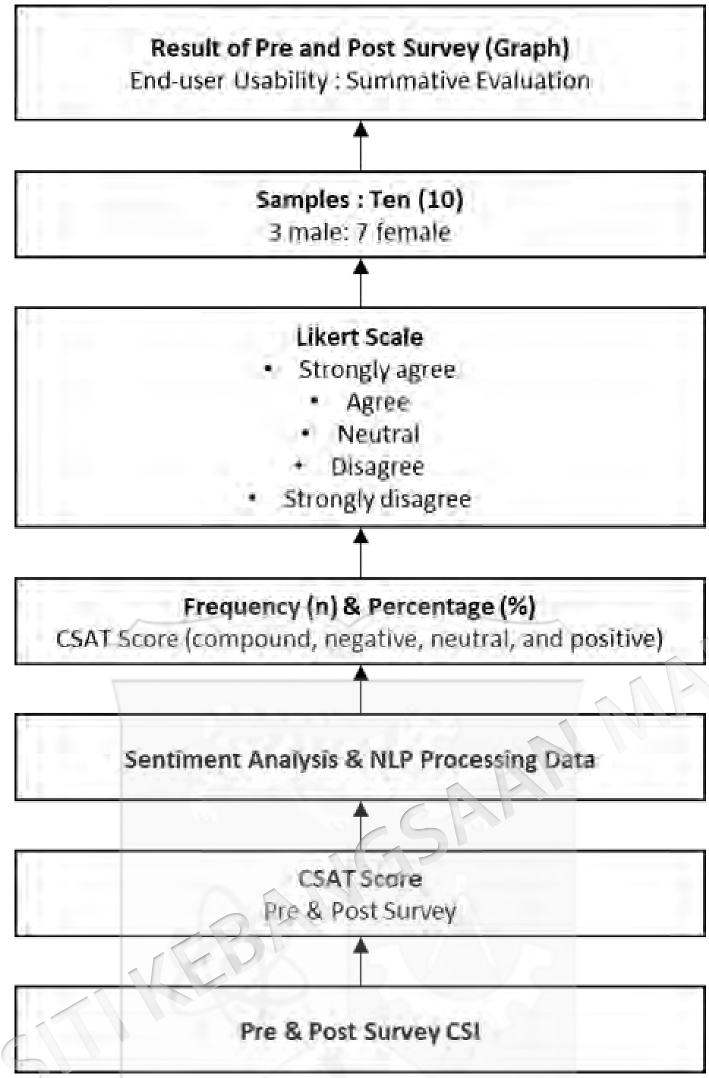


Figure 5.20 Holistic Attribute Analysis Model: Users' Level of Satisfaction between Chatbot & New Improved Chatbot (HAAM-ULS-C &NC)

A total of ten (10) people participated in the pre-survey. The analysis measured the sample's sentiment level using the CSAT score. The score ranges from one to five which are from very unsatisfied, unsatisfied, neutral, satisfied, and very satisfied. Another measurement for sentiment used in text was through natural language processing which ranged are compound, negative, neutral, and positive statement. Figure 5.21 shows the results of the findings obtained from the analysis of the pre-survey conducted.

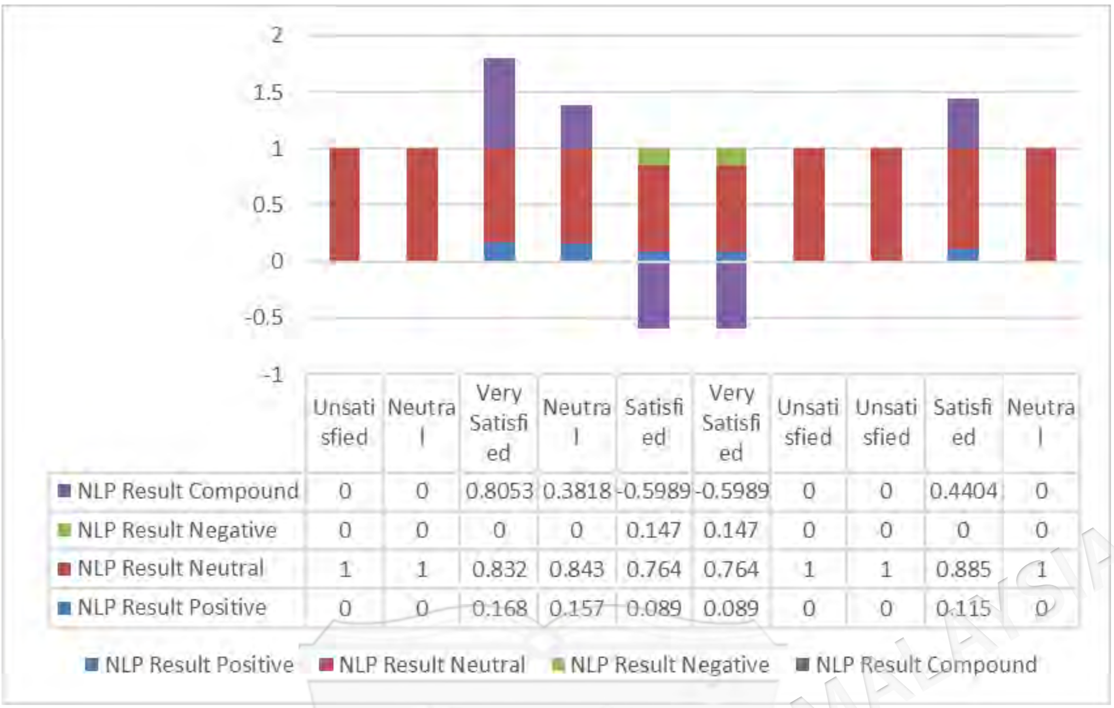


Figure 5.21 Pre-survey Sentiment Analysis & CSAT Score

Based on Figure 5.21, it can be observed that 40% of the sample were "satisfied" and "very satisfied" while 30% were neutral. It can be inferred that majority (70%) were "unsatisfied" or "very unsatisfied" with the chatbot. This also clearly observed in Table 5.22 that shows the CSAT score recorded during pre-survey.

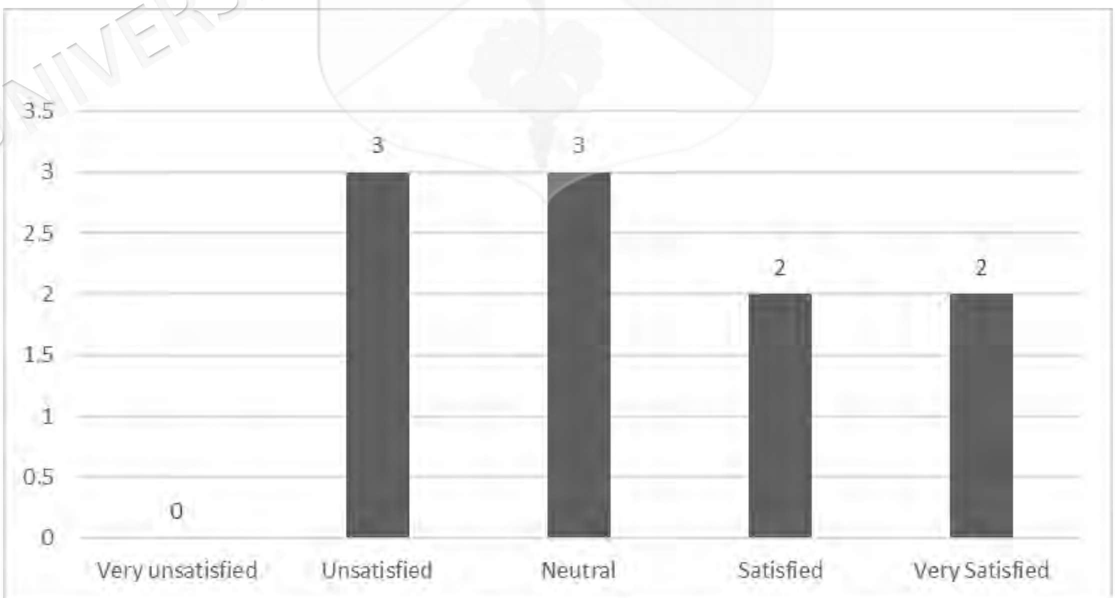
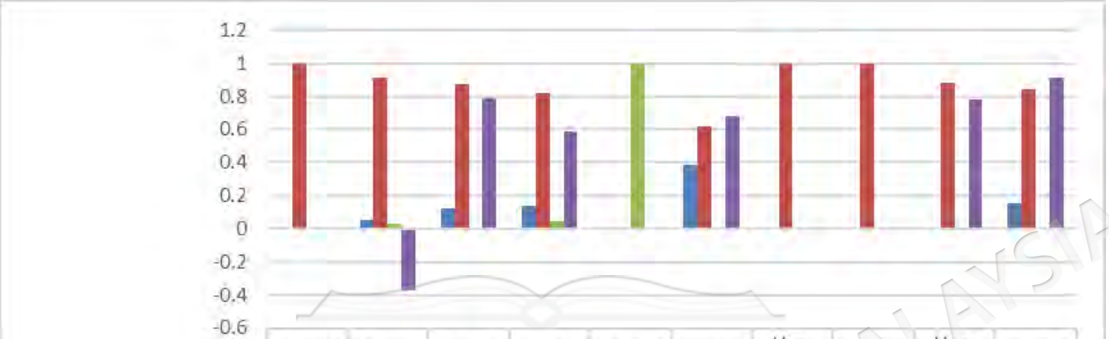


Figure 5.22 CSAT Score for pre-survey

The feedback obtained from the pre-survey resulted in post-survey of which resulted in the integration of new attributes for the improved version of chatbot. A total of ten (10) samples participated in the post survey. The same set of instruments were given to each sample. Figure 5.23 recorded the sentiment analysis as well as the natural language processing data for the post survey in the new improved chatbot.



	Neutral 1	Neutral 2	Satisfied 3	Satisfied 4	Neutral 5	Satisfied 6	Very Satisfied 7	Satisfied 8	Very Satisfied 9	Neutral 10
NLP result Positive	0	0.053	0.124	0.137	0	0.384	0	0	0.0115	0.153
NLP result Neutral	1	0.914	0.876	0.82	0	0.616	1	1	0.885	0.847
NLP result Negative	0	0.033	0	0.043	1	0	0	0	0	0
NLP result Compound	0	-0.3736	0.7882	0.5859	0	0.6808	0	0	0.782	0.9101

NLP result Positive   NLP result Neutral   NLP result Negative   NLP result Compound

Figure 5.23 Post-survey sentiment analysis and CSAT score on New Chatbot

Findings of the post-survey conducted on the new improved chatbot showed that 50% of the samples were "satisfied" and "very satisfied", whilst 50% were neutral. This can be clearly observed from the Figure 5.24. This is also clearly seen from the comparison of CSAT score of the pre- and post-survey indicated from the line graph shown in Figure 5.25.

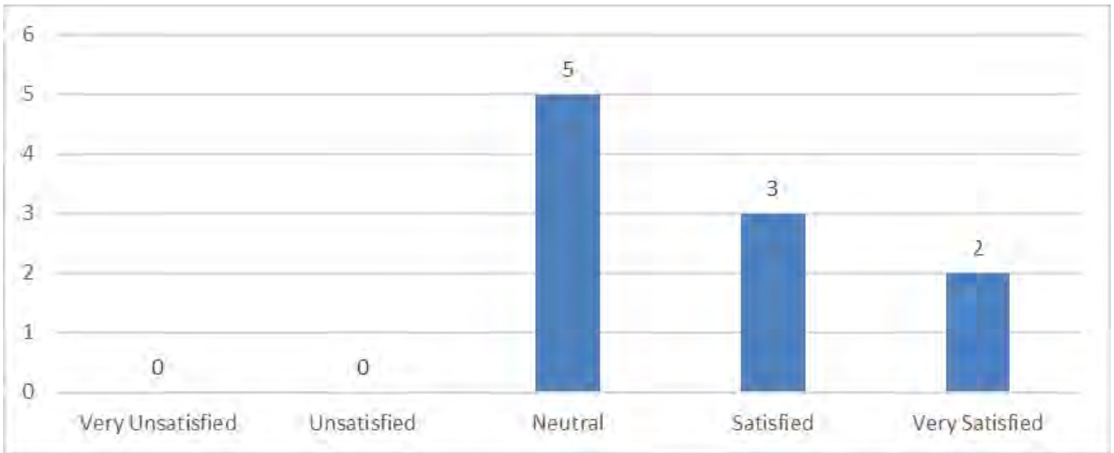


Figure 5.24 CSAT Score post-survey

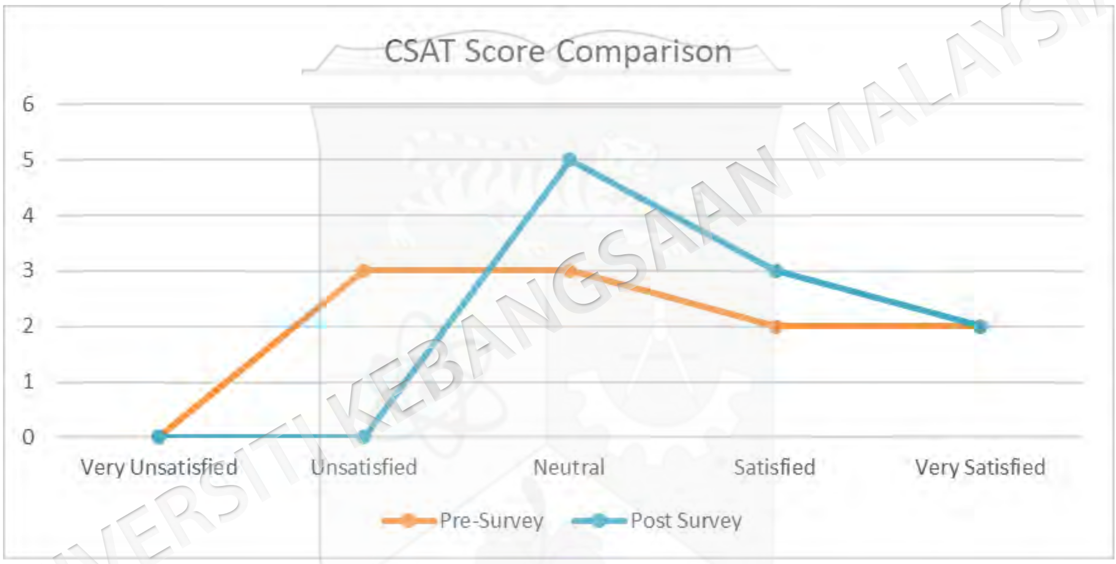


Figure 5.25 CSAT Score comparison between pre-development and post-development

5.9 HYPOTHESES TESTING

There are two (2) hypotheses that were tested in this research. The following section will discuss the findings for each hypothesis. Table 5.7 below is the indicator of the Pearson Correlation Coefficient (r) value which act as guidance to answer the hypothesis one and two (H01 & H02).

H01: There is no significant difference between quality features and customer satisfaction level.

H02: There is no correlation between positive sentiment analysis and customer satisfaction level.

Table 5.7 Pearson correlation coefficient value

Pearson Correlation Coefficient (r) value	**Strength**
.91 until 1.00 or -.91 until -1.00	Very strong
.71 until .90 or -.71 until -.90	Strong
.51 until .70 or -.51 until -.70	Moderate
.31 until .50 or -.31 until -.50	Weak
.01 until .30 or -.10 until -.30	Very weak
.00	No correlation

H01: There is no significant difference between quality features and customer satisfaction level.

Table 5.8 Pearson correlation (H01)

		Functionality	Efficiency	Technical Satisfaction	Humanity	Effectiveness	Ethics	Customer Satisfaction
Functionality	Pearson Correlation	1	.588	.731*	.871**	.913**	.842**	**.748***
	Sig. (2-tailed)		.074	.016	.001	.000	.002	**.013**
	N	10	10	10	10	10	10	**10**
Efficiency	Pearson Correlation	.588	1	.597	.706	.569	.458	**.615**
	Sig. (2-tailed)	.074		.068	.022	.086	.183	**.058**
	N	10	10	10	10	10	10	**10**
Technical Satisfaction	Pearson Correlation	.731*	.597	1	.853	.687	.700	**.791****
	Sig. (2-tailed)	.016	0.68		.002	.028	.024	**.006**
	N	10	10	10	10	10	10	**10**
Humanity	Pearson Correlation	.871**	.706	.853*	1	.841**	.867**	**.671***
	Sig. (2-tailed)	.001	.022	.002		.002	.001	**.034**
	N	10	10	10	10	10	10	**10**
Effectiveness	Pearson Correlation	.913**	.569	.687	.841**	1	.884**	**.740***
	Sig. (2-tailed)	.000	.086	.028	.002		.001	**.014**
	N	10	10	10	10	10	10	**10**
Ethics	Pearson Correlation	.842**	.458	.700*	.867**	.884**	1	**.493**
	Sig. (2-tailed)	.002	.183	.024	.001	.001		**.147**
	N	10	10	10	10	10	10	**10**
Customer Satisfaction	Pearson Correlation	.748*	.615	.791**	.671*	.740*	.493	1
	Sig. (2-tailed)	.013	.058	.006	.034	.014	.147	
	N	10	10	10	10	10	10	10

Table 5.8 shows the outcomes form the Pearson correlation testing between attributes of the chatbot and customer satisfaction level. Based on the results shown, there is a strong positive correlation (r = .748) and statistically significant with p-value less than 0.05 (p = .013). Thus, there is significant difference between functionality and customer satisfaction level. As for efficiency, the results show moderate positive correlation (r = .615) and not statistically significant (p = .058). Thus, there is no significant difference between efficiency and customer satisfaction level. For technical satisfaction, it is recorded that there is strong positive correlation (r= .791) and statistically significant (p=.006). Thus, there is significant difference between technical satisfaction and customer satisfaction level. Next, for humanity, there is moderate correlation (r=.671) and statistically significant (p=.034). Thus, there is significant difference between humanity and customer satisfaction level. For effectiveness, there is strong correlation (p=.740) and statistically significant (r=.014). This shows there is significant relationship between effectiveness and customer satisfaction level. The last attribute tested was ethics and there is very weak correlation (p=.493) and not statistically significant (r=.147). thus, there is no significant difference between ethics and customer satisfaction level.

Table 5.9 Pearson correlation (attribute and customer satisfaction)

		Attribute_Mean	Customer Satisfaction
Attribute_Mean	Pearson Correlation	1	.7689
	Sig. (2-tailed)		.0093
	N	10	10
Customer Satisfaction	Pearson Correlation	.7689	1
	Sig. (2-tailed)	.0093	
	N	10	10

Table 5.9 recorded the Pearson Correlation testing outcomes between mean of all attributes tested against customer satisfaction level. Based on the results, there is strong correlation (r = .7689) with statistically significant relationship between these two (2) variables (p = .0093). Thus, it successfully rejects the null hypothesis (HO1). This indicates that good/quality attributes contribute to higher customer satisfaction level.

H02: There is no correlation between positive sentiment analysis and customer satisfaction level.

Based on testing done using Pearson Correlation testing, the following table shows the result of two (2) phase which is between positive sentiment analysis and customer satisfaction level during pre and post development phase.

Table 5.10 Pearson correlation (pre-development)

		Sentiment_Pre-Positive	Satisfaction_Pre
Sentiment_Pre_Positive	Pearson Correlation	1	**-.020**
	Sig. (2-tailed)		**.956**
	N	10	**10**
Satisfaction_Pre	Pearson Correlation	**-.020**	1
	Sig. (2-tailed)	**.956**	
	N	**10**	10

Table 5.10 recorded the result from Pearson Correlation testing toward positive sentiment analysis towards customer satisfaction level during pre-development phase. The outcomes were a very weak (r=-.020) and non-significant correlation (p=.956) between positive sentiment analysis and customers satisfaction level. This suggest that any relationship between these two variables during pre-development phase is likely due to chance.

Table 5.11 Pearson correlation (post development)

		Sentiment_Post-Positive	Satisfaction_Post
Sentiment_Post_Positive	Pearson Correlation	1	**.274**
	Sig. (2-tailed)		**.443**
	N	10	**10**
Satisfaction_Post	Pearson Correlation	**.274**	1
	Sig. (2-tailed)	**.443**	
	N	**10**	10

Table 5.11 shows the result from Pearson Correlation testing towards positive sentiment analysis towards customer satisfaction level during post development phase. From the result, it is found that there is a weak positive correlation (r=.274), however it is not statistically significant (p=.443). Thus, it is failed to reject the null hypothesis

(HO2). This indicate that positive sentiment analysis does not really contribute to customer satisfaction level.

5.10 CONCLUSION

The findings of this study were based on the four (4) main objectives: (i) to determine the attributes for a good quality chatbot; (ii) to develop the text analyzer technique; (iii) to explore and develop the sentiment analysis for the chatbot;(iv) to validate the user's or customer's level of satisfaction on the new improved chatbot. For all these objectives, the necessity to develop a methodology called the Holistic Evolutionary Iterative Chatbot Development Life Cycle (HEI-HC-LC) to ensure the successful development of the New User Humanity-Like Chatbot (NURUL-C).

Therefore, the findings to the development and verification of the quality attributes or the attribute for a good quality chatbot was done through the dimensional or category of attributes together with the sub-dimension or sub-categories that were selected through past research works through literature review and also through the pre-survey and post survey; the sentiment analysis and the customer satisfaction index (CSI) that was conducted using various instruments: questionnaire and observational checklist were designed for the purpose. A panel of six (6) experts in software development, HCI, NLP, data mining, and banking validated and verified the instrument.

Additional discoveries involved the text analyzer, which was employed to scrutinise the vocabulary, text, and interactions between the chatbot and its users or customers. Various approaches and techniques were used such as the word tokenisation, text normalisation, stemming, and lemmatisation.

Findings on the Pre and Post test showed interesting results which indicated that the specified selected quality attributes improved the sentiment analysis as well as the customer satisfaction index. An interesting element in the attribute category is the energy efficiency which is in line with the sustainability and energy saving that is so important at this time of energy transition where nation is trying to improve the sustainability of the environment to net zero carbon by 2050 or earlier. This element can be seen in term of time taken for the chatbot to response, and size taken for chatbot to be place on-premises.

The other finding is the improved new chatbot prototype developed through the selected quality attributes called New User Humanity-like Chatbot (NURUL-C). Result of the end-user usability evaluation (summative) evaluation shows positive results, although much can still be improved. Finally, the other significant finding is the result of the Customer Analysis, and the Customer Satisfaction Index (CSI) based on the CSAT score.

CHAPTER VI

IMPLICATIONS AND CONCLUSION

6.1 INTRODUCTION

The research used a mixed-mode methodology to design a high-quality chatbot by identifying quality attributes, sentiment analysis techniques, and the use of Customer Satisfaction Index (CSI) for usability. The study had focused on the banking domain, particularly the retail aspect of the banking system that involves customers with the daily banking services requirement. This chapter discusses on the research summary; the main findings in relation to the objectives of the research; discusses on the design and development of the new improved chatbot namely: New User Humanity-like Chatbot (NURUL-C) and as well as the evaluation of NURUL-C based on the sentiment analysis and the Customer Satisfaction Index (CSI); implications of the research findings; and recommendations for future research that can be conducted.

6.2 RESEARCH SUMMARY BASED ON RESEARCH RESULTS

This research involves summaries of the findings based on the four (4) main objectives as follows:

i. Determining the attribute for a good quality chatbot. This was based on the theoretical framework, research conceptual framework, Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle (HEI-HC-LC), and design workflow. Through all these frameworks and modules, it had directly and indirectly helped determined the selection of the attributes systematically. Attributes of chatbots from past research were studied, scrutinised, and selected through the processes involved; verified by the panel of experts and tested through formative and summative evaluation (user usability evaluation).

Eventually, six (6) categories of attributes were selected and considered the quality attributes that would result in the creation of a good quality chatbot namely: Functionality Adequacy (which included the function of visual); Efficiency (which included energy efficiency); Technical Adequacy; Humanity (which included elements of realism on the part of the chatbot); Effectiveness (in terms of the natural language used and its accuracy); and Ethics (which included understanding of users' ethics requirements).

ii. Design and development of a new improved chatbot. This aspect of the study involved the development of the chatbot development methodology that integrated three (3) phases of the development, which later verified the attributes and then conduct evaluation through sentiment analysis and customer satisfaction index called the Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle (HEI-HC-LC).

iii. Developing of the text analyser technique for the new user humanity-like chatbot (NURUL-C) using stemming and lemmatisation approaches.

iv. Conduct the sentiment analysis (Pre and Post Analysis). A sentiment analysis was conducted to ascertain user's input on the current available chatbot, and the potential quality attributes based on their effectiveness and usefulness (through 5-Likert Scale) and recorded statements. The evaluation was also conducted on the Customer Satisfaction Index (CSI) based on CSAT Score.

6.2.1 Determining the Attributes for a Good Quality Chatbot

The objective of determining the attributes for a good quality chatbot was to identify quality attributes to be selected and integrated with the design and development of the new improved chatbot: the New User Humanity-like Chatbot (NURUL-C). Thus, identifying and determining the attributes would provide plausible solutions to help customers run their daily routine banking chores through the services provided by the chatbot. The six (6) categories of attributes selected and used in a pre, and post survey were administered to the users. Before they were used and integrated in the chatbot, these categories of attributes were verified by the Panel of Experts, who had specific knowledge on the banking domain, customer information seeking behaviours, the Human-Computer Interaction (HCI), Data Mining, and Natural Language Processing

(NLP). They helped to verify the categories and the dimensions as well as elements based on their categories. The categories were also based on works conducted by researchers earlier (Radziwill & Benton 2017; Kaleem et al. 2016) and adapted to meet the technology and users of the current landscape. ISO 9241 which supporting the concept of usability emphasises that "the effectiveness, efficiency, and satisfaction with which specified users achieve specified goals in particular environments."

6.2.2 Design and Development of a New Improved Chatbot

The design and development of a new improved chatbot called the New User Humanity-like Chatbot (NURUL-C), involved the development of the overall development methodology called the Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle (HEI-HC-LC), which involved the pre-development; design, development, and implementation; and the evaluation phase. The pre-development phase encompassed preliminary study on attributes underpinned by theories and previous studies. The design, development, and implementation phase encompassed the determination and selection of quality attributes that were selected, verified, and integrated into the new improved chatbot, NURUL-C.

6.2.3 Developing the Text Analyser Technique

The text analyser technique adopted in the new chatbot, NURUL-C was stemming and lemmatisation interpretation in Natural Language Processing (NLP). In NLP, there are two (2) divisions: the Natural Language Processing generation and natural language understanding. The former, involves producing meaningful phrases and sentences in the form of natural language in which text planning, sentence planning, and text realisation occurs. The latter, involves better understanding of the text which encompasses mapping the given input in natural language into useful representation for better analysis of the language. In relation to this, stemming, a process which produces variant of a root or basic word. In other words, it reduces a base word to its stem word. Thus, stemming was used in the text analyser for NURUL-C, to shorten the look-up and normalise the sentences for a better understanding. In addition to stemming, lemmatisation was also applied in the text analyser of NURUL-C. Lemmatisation is the process of assembling the inflected parts of a word, such that they can be recognised as a single element, hence called the word's lemma or its vocabulary form. This process is like stemming but it adds meaning to particular words; it connects text with alike meanings to a simple word. In actual situation, although performing stemming is easier that than the process of lemmatisation, but the latter makes more meaning and understanding, thus proves to be the ultimate choice. The use of the two (2) approaches as the text analyser technique applied in NURUL-C, had helped make the conversations between the user and the chatbot, more humane-like, realistically human, and understandable.

6.2.4 Sentiment Analysis and Customer Satisfaction Index (CSI)

Sentiment Analysis were conducted at the pre-development as well as the evaluation phase. The pre-development phase involved preliminary survey and pre-survey (formative evaluation). The Evaluation phase involved the Pilot Survey, and the post-surveyor end-user usability survey (summative evaluation). There were conducted through observations and test conducted by users using observational checklists as instruments; and survey administered to users after a hands-on experience with the chatbot through questionnaires as the instrument. The preliminary and the pre-survey conducted showed that current chatbots were not positively perceived as they did not pose attributes that could make the users feel at ease as the chatbots were considered not 'humanity' friendly. The selected 'quality' attributes were verified by the six (6) Panel of Experts and then integrated into the new user's humanity-like chatbot (NURUL-C). An end user usability evaluation (summative evaluation) was conducted as a post-survey on the improved NURUL-C. Results showed a significant positive change to the effectiveness of the chatbot.

Therefore, findings from the formative and summative evaluations shows that users' or customers' sentiment analysis had improved, as more positive sentiment was detected through the statements that were recorded from the post-survey administered to the users or customers. The 'quality' attributes that were integrated into the improved chatbot, NURUL-C, had increased the score of the effectiveness of the service provided through the improved chatbot. Most interesting was the introduction of the energy efficiency elements, the visual elements, and the humanity dimension and its elements.

These and the other attributes had made the chatbot more realistic with its greetings, 'smiling robot', and human-like response.

All the sentiment analysis and end-user usability evaluation were analysed based on the Holistic Attribute Analysis Models (HAAMC-I-HAAM-C12) that can be observed in Figure 5.1; Figure 5.8 – 5.12. The results of the study indicate that the attributes that were integrated into the design and development of a chatbot, played a significant role. In this study it was clear that the integration of the six (6) quality categories of attributes: its dimensions and elements has resulted in the better acceptance of the users, as the evidence showed that the user or Customer Satisfaction Index (CSI) had improved.

6.3 IMPLICATIONS OF RESEARCH

The implication of the research can be summarised into four (4) aspects:

i. Implications of research on the determination of the attributes for a good quality chatbot based on the six (6) categories selected.
ii. Implications of research on the design and development of the new improved chatbot, NURUL-C.
iii. Implications on the text analyser technique adopted in the new improved NURUL-C chatbot.
iv. Implications of the Sentiment Analysis and Customer Satisfaction Index (CSI).

a. Implications of Research Results on the Determination of the Attribute for a Good Quantity Chatbot

The implications of research results in the identification and determination of the attributes for a good quality chatbot is very significant. This is because, in many of the past research that were conducted, it was found that the attributes integrated into the chatbots would have a significant contribution to the chatbot. A carefully selected 'quality' attribute would result in a good quality chatbot that would be able to satisfy customer needs better. Similar results were found in the current study conducted. The six (6) selected categories of attributes (their dimensions and elements) had resulted in the new improved chatbot, NURUL-C securing more positive result in comparison to the available current chatbots used by the current banking settings.

Thus, this means that in the banking system, application of chatbots of the future can make use of the 'quality' attributes determined from this study. Software developers that are developing future chatbots can used these attributes and add any new attributes that are more 'humanity-like' that was the preference of users or customers in Malaysia. These attributes would have great implications and impact to any new and future chatbots that will be developed for the banking environment in Malaysia. They need not be cosmetically acceptable in appearance, but in the banking domain they need to be humanistic-ally accepted.

b. Implications on Design and Development of a New Improved Chatbot

The implications of the research results on the design and development of the improved chatbot, NURUL-C, has significant implication on the design and development of future chatbot applications for the banking domain or setting. The development of a good quality chatbot should not only be based on 'quality' attributes but also on a systematic process of application design and development so that the use of the application in this case the chatbot, can have a positive impact on the users and customers. Thus, the application engineering process based on a sound development life cycle methodology is crucial.

The development of the Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle (HEI-HC-LC) had important implication on service application development specifically for banking customers. This user-centered life cycle approach focused on to understand the humanistic and humanity-like aspects of the banking customers' needs, which can be a useful guide for banking systems or applications (chatbot) developers in Malaysia.

The use of visual elements and the energy efficient elements as well as the humanity dimension and its elements to be integrated into the design of the new improved chatbot, NURUL-C, also had implications for the chatbot developers in the service sector domain. Based on this current study, the use of these and other attributes, which holistically contributes to a more 'humanity-like' chatbot was an interesting 'eye-opener' to the banking customers and can be a verified guide to future chatbot developers in the banking and other service sector domains.

c. Implications on the Text Analyser Technique Adopted in the Improved Chatbot, NURUL-C

The implications on the text analyser technique adopted in the improved chatbot, NURUL-C, had an impact in the customers of the banking domain as well as other service sector domains. The use of stemming that can shorten text of words from the root word; and lemmatisation that makes the text of words to be more meaningful and understandable, meant that conversations can take place in a more meaningful manner between the chatbot and the customers. The use of Natural Language Processing (NLP) approach also meant the language used is the language that normally used by the customers and this very humanistic and humanity-like communication appeals to the customers.

d. Implications of the Sentiment Analysis and Customer Satisfaction Index (CSI)

The research results in the sentiment analysis and Customer Satisfaction Index (CSI) based on the formative and summative evaluation administered to the users or customers using the observational checklist and questionnaires as the instruments, have significant impact on the effectiveness and usefulness of the chatbot in providing a service to the banking domain. The result of the formative evaluation indicated that there was a need for an improved chatbot with quality attributes; and the results of the summative evaluation showed that customers were more satisfied with the improved chatbot (NURUL-C) that was integrated with selected quality attributes. Therefore, it is imperative that before a chatbot designer and developer begin the design and development of a chatbot, they will have to understand well their potential user, particularly on the 'humanity' aspect, their needs and information seeking behaviour.

e. Implications for the Banking Sector

Research result of the current study has great implications for the banking sector. The banking sector needs to realise that the services that they provide through the tool such as a chatbot need to carefully be designed and developed. The design must include various factors such as the best and most recent information or knowledge that need to be provided by the chatbot to the customers: the quality 'humanity-like' attribute; the needs of the customers; the customers information seeking behaviour and their Customer Satisfaction Index. Taking into consideration on a sound design and development approach would result in a more effective and usable chatbot that can benefit both the banks and their customers.

6.4 RECOMMENDATIONS FOR FUTURE RESEARCH

Based on summary of the findings of the study discussed, several recommendations for future research can be conducted to meet the theoretical and practical gaps of the study. There are the areas that includes software engineering, AI user-centered chatbot design, Human-Computer Interface (HCI), Sentiment Analysis, Data Mining, and NLP in the banking domain. Among the recommendation for future research cover fields as follows:

a. Field of Software Engineering

Further research can be conducted in future based on the current research. The field of software engineering involve research in the field of system, application including chatbot development of a holistic life-cycle development model, such as the one developed in the current study called Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle (HEI-HC-LC) of the New User-Centered Humanity-like Chatbot (NURUL-C), for users in the services sector, particularly the banking sector. The model can be further researched through more detailed development of process, where comments from users or customers can be enhanced and verifications be measured through a suitable statistical analysis, to improve the model. The verification of the attributes too can be further improved from the current process, which is approval by a panel of experts. It can still be conducted by the Panel of Experts but conducted more formally and in a structured manner through a three(3)-round Delphi approach. This would be more reliable and convincing. The user-centered 'humanity-like' approach in the design and development of this model could be improved with new enhanced elements to be incorporated, if necessary, new principles can be built through additional dimensions and elements for the chatbot. This can contribute to both the theoretical and practical gaps of the Software Engineering or Application Engineering field.

b. Field of AI in User-Centered Humanity-like Chatbot Design

The field of 'User-centered Humanity-like' chatbot falls within the field of Artificial Intelligence (AI). In the current study, the user centered humanity-like chatbot was designed based on the user-centered design. However, future user-centered chatbots that is built within the 'humanity-like' framework, needs to integrate AI algorithm that will make the chatbot more humanity-like in terms of the conversational experience and interface design. In future work in this area, the use of NLP and ML needs to be integrated so that the chatbot can decipher the customer's query and decide on the responds more accurately and satisfying. This can contribute to both the theoretical and practical field of AI in user-centered humanity-like design of chatbots.

c. Field of Human-Computer Interaction (HCI)

Future research can be conducted in the field of Human-Computer Interaction (HCI) specifically in the design, development, and strategies of customer information seeking behaviour for the service sector, specifically the banking sector. This study used the user-centered interaction approach through the visual and text conversational approach. This can be improved by integrating AI with media to make it more audio and visual, conversational approach which will be more 'humanity-like' and more personally meaningful. This can then contribute to both the theoretical and practical gaps of the HCI field.

d. Field of Sentiment Analysis

Further future research can be conducted in the field of Sentiment Analysis. In this study, sentiment analysis conducted was based on opinion mining (aspect in data mining) and Natural Language Processing (NLP) technique performed on textual data, using the text analyser technique. Although it is sufficient in monitoring the service sentiment based on the customer feedback, it could be further improved by using the many other techniques that could be applied through the integration of AI techniques that would be able to better understand specific customer's needs, thus improved their satisfaction index. This can contribute to the theoretical and practical gaps of the sentiment analysis field.

e. Field of Banking and Finance

Further research can also be conducted on the field of banking and finance in the context of service product knowledge that need to be integrated into the chatbot application. What are the new potential products of not just retail but corporate banking and finance including Islamic Fintech (Islamic digital financing). This can then contribute to both the theoretical and practical gaps of the banking domain.

6.5 CONCLUSION

The research has successfully fulfilled the research objectives, research question and hypotheses. The research had also solved the problem set earlier for the current research that covered aspects on: the determination of the attributes that would be suitable for a good quality chatbot; the design and development of a new improved chatbot, namely New User Humanity-like Chatbot (NURUL-C); development of the text analysis; and finally, evaluation of the chatbot, through sentiment analysis and Customer Satisfaction Index (CSI) based on the CSAT score. The evaluation was conducted both at the formative and summative levels.

The determination of the six (6) selected quality attributes were conducted based on previous chatbots developed and in past research conducted reviewed in the literature. The final selected attributes were verified by six (6) panel members of experts in the multidisciplinary fields involved in the research. The attributes were integrated into the design of the new improved chatbot, NURUL-C and were assessed with the customers using the questionnaire as the instrument. The six (6) attributes can be the initial start for designers and developers intending to build future service chatbots, especially for the banking and finance (including Islamic Banking Finance as well as Islamic Digital Finance or FinTech) sector.

The Holistic Evolutionary Iterative Humanistic Chatbot Development Life Cycle (HEI-HC-LC) model for designing and developing the New User Centered Humanity-like Chatbot (NURUL-C), for users or customers in the service sector (particularly the banking sector), has shown to be a successful approach. The approach which encompasses three (3) fundamental phases; Phase I: Pre-development (which involved the functional analysis, Preliminary study, also known as Formative Analysis on attributes). It is at the second phase (II) too, that the quality attributes were selected and verified; Phase III: Evaluation (Summative Analysis or Evaluation); meant that the design and development of the new improved chatbot underwent an intensive and rigorous process of design and development to ensure that the improved chatbot meet the needs of the users or customers. The fact that the development methodology model adopted the standard software development life-cycle, with specific requirements for chatbot design and development, made it a very systematic and comprehensive approach that not only took into consideration of the physical aspects of the chatbot, but also the humanistic as well as the quality assurance aspects of the chatbot development.

The research result of the evaluation conducted, based on the formative as well as the summative evaluation where all the details and findings are discussed in Chapter V, showed that the quality attributes are crucial in the design and development of a good quality chatbot, that would meet the needs of the users or customers. In conclusion, it can be confirmed that the research result has achieved the research objectives and have successfully contributed to the existing theoretical and practical gap of the related fields of knowledge applied in this study. The current research has managed to integrate the multidisciplinary fields into the domains of chatbot application technology in the service sector of banking effectively and meaningfully.

REFERENCES

Adamopoulou, E., & Moussiades, L. 2020. Chatbots: History, technology, and applications. *Machine Learning with Applications* 2: 100006.

Aggarwal, K., Mijwil, M. M., Al-Mistarehi, A. H., Alomari, S., Gök, M., Alaabdin, A. M. Z. & Abdulrhman, S. H. 2022. Has the future started? The current growth of artificial intelligence, machine learning, and deep learning. *Iraqi Journal for Computer Science and Mathematics* 3(1): 115-123.

Aishwarya, B., Vijay, G. & Sachin, D. 2017. Chatbot for business organization: Utility, processes and significance. *International Journal of Scientific Research in Computer Science, Engineering, and Information Technology* 2(6): 872-881.

Akram, A., Mousumi, S., Tamasree, B. 2019. Evaluation of descriptive answer sheet using artificial intelligence. *International Journal of Engineering Sciences and Research Technology* 8(4): 184-186.

Annett, M. & Kondrak, G. 2008. A comparison of sentiment analysis techniques: Polarizing movie blogs. *Conference of the Canadian Society for Computational Studies of Intelligence*, pp. 25-35.

Aprilinda, Y., Martavia, T., Erlangga,E., Afandi, F., N. & Rizal, U. 2022. Chatbot menggunakan natural language processing untuk Pembelajaran Bahasa Inggeris berbasis android. *EXPERT: Jurnal Manajemen Sistem Informasi dan Teknologi* 12(1): 23.

Aronson, J.E., Liang, T.P. & MacCarthy, R.V. 2005. *Decision Support Systems and Intelligent* Systems. Vol. 4. Upper Saddle River, NJ, USA: Pearson Prentice-Hall.

Arpowosaye, T. M. & Bakare, O. D. 2022. Assessing information-seeking behavior of visually impaired students using Wilson’s model of information behavior in federal college of education (special), Oyo. *Library Philosophy and Practice (e-Journal)* 7202.

Baccianella, S., Esuli, A. & Sebastiani, F. 2010. SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Lrec* 10: 2200-2204.

Balve, M. 2015. Messaging apps are now bigger than social networks. http://uk. businessinsider. com/the-messaging-app-report-2015-11 [3 July 2025].

Balvir, S., Sahu, S. & Rohankar, J. 2021. Approaches and applications of sentiment analysis on users data. *Journal of University of Shanghai for Science and Technology* 23(6): 1761-1767.

Bayu, S. & Ferry, W. W. 2016. Chatbot using a knowledge in databased. *The International Conference on Intelligence Systems Modelling and Simulation*, pp. 72-77.

Beale, R. & Creed, C. 2009. Affective interaction: How emotional agents affect users. *International Journal of Human-Computer Studies* 67(9): 755-776.

Ben, W. & Karaa, A. 2013. A new stemmer to improve information retrieval. *International Journal of Network Security & Its Applications* 5(4): 143-154.

Bhatia, P., Ji, Y. & Eisenstein., J. 2015. Better document-level sentiment analysis from rst discourse parsing. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 2212-2218.

Birjali, M., Kasri, M. & Beni-Hssane, A. 2021. A comprehensive survey on sentiment analysis: Approaches, challenges, and trends. *Knowledge-based Systems* 226: 107134.

Birkett, A. 2023. What is Customer Satisfaction Score (CSAT) and how to measure it? https://blog.hubspot.com/service/customer-satisfaction-score#what-is [20 November 2023].

Biswas, P., Neogi, S. G., Arockiam, D. & Mitra. A. 2024, March. The importance of multi-lingual sentiment analysis in digital marketing. *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Direction)*, 1-15.

Boltayevich, E. B., Mirdjonovna, H. S. & Ilxomovna, A. X. 2023. Methods for creating a morphological analyzer. In. Zaynidinov, H., Singh, M., Tiwary, U.S., Singh, D. (eds). *Intelligent Human Computer Interaction, IHICI2022: Lecture Notes in Computer Science*, V0l.13741. Cham: Springer.

Borah, B., Pathak, D., Sarmah, P., Som, B. & Nandi, S. 2019. Survey of text-based chatbot in perspective of recent technologies. In *Computational Intelligence, Communications, and Business Analytics: Second International Conference CICBA 2018*. Pp. 84-96.

Bradeško, L. & Mladenić, D. 2012. A survey of chatbot systems through a Loebner prize competition. *Proceedings of Slovenian Language Technologies Society Eight Conference of Language Technologies*, pp. 34-37.

Brady, S. R. 2015. Utilizing and adapting the Delphi method for use in qualitative research. *International Journal of Qualitative Methods* 14(5): 1-6.

Brandtzaeg, P. B., & Folstad, A. 2017. November. Why people use chatbots. *International Conference on Internet Science*, pp. 377-392.

Brooke, J. 1996. SUS: A quick and dirty usability scale. In. Jordan, P. W., Thomas, B., Weerdmeeser, B. A. & McClelland, I. L. (eds.). *Usability Evaluation in Industry*, 189-194. London: Taylor & Francis.

Candela, E. 2018. Consumer's perception and attitude towards chatbot adoption. A focus on the Italian market. Master's Thesis, Aalborg University, Dernmark.

Candello, H., Vasconcelos, M. & Pinhanez, C. 2017. Evaluating the conversation flow and content quality of a multi-bot conversational system. *Proceeding of the 2017 CHI Conference Extended Abstract on Human Factors in Computing System*, pp 60-61.

Chatbots Magazine. 2019. Chatbot report 2019: Global trends and analysis. 19 April 2019. https://chatbotsmagazine.com/chatbot-report-2019-global-trends-and-analysis-a487afec05b [26 April 2022].

Chiravirakul, P. & Payne, S. J. 2014. Choice overload in search engine use?. *Proceedings of the SIGCHI Conference of Human Factors in Computing Systems*, pp. 1285-1294.

Cohen, D. & Lane, I. R. 2016. An oral exam for measuring a dialog system's capabilities. *AAAI Conference on Artificial Intelligence*, pp. 835-841.

Cohen, P. R. 2005. If not Turing's Test, then what?. *AI Magazine* 26(4): 61-67.

Collumb, A., Costea, C., Joyeux, D., Hasan, O. & Brunie, L. 2014. A study and comparison of sentiment analysis methods for reputation evaluation. https://liris.cnrs.fr/Documents/Liris-6508.pdf [3 July 2025].

Coniam, D. 2014. The linguistic accuracy of chatbots: Usability from an ESL perspective. *Text Talk* 34(5): 545-567.

Crescenzi, A., Kelly, D. & Azzopardi, L. 2015, August. Time pressure and system delays in information search. *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 767-770.

Dale, R. 2016. The return of chatbots. *Natural Language Engineering* 22(5): 811-817.

Dalkey, N. & Helmer, O. 1963. An experimental application of the Delphi method to the use of experts. *Management Science* 9(3): 458-467.

Damayanti, L. & Lhaksmana, K. M. 2024. Sentiment analysis of the 2024 Indonesia presidential election on Twitter. *SinkrOn* 8(2): 938–946.

Delacroix, L. 2007. *Longman Advanced American Dictionary*. Edinburgh: Pearson Education.

Devika, M. D. Sunitha, C., & Ganesh, A. 2016. Sentiment analysis: A comparative study on different approaches. *Procedia Computer Science* 87: 44-49.

Divya, K., Sidhartha, B. S., Niveditha, N. M. & Divya, B. M. 2020. An interpretation of lemmatization and stemming in natural language processing. *Journal of University of Shanghai for Science and Technology* 22(10): 350-357.

Dix, A., Finlay, J., Abowd, G. D. & Beale, R. 2004. *Human-Computer Interaction*. 3rd Edition. Edinburgh: Pearson Education.

Donohoe, H. M. & Needham, R. D. 2009. Moving best practice forward: Delphi characteristics, advantages, potential problems, and solutions. *International Journal of Tourism Research* 11(5): 415-437.

Duijst, D. 2018. Can we improve the user experience of chatbots with personalisation? Master's Thesis, University of Amsterdam, Netherlands.

Ebrahimi, M., Yazdavar, A. H. & Sheth, A. 2017. Challenges of sentiment analysis for dynamic events. *IEEE Intelligent Systems,* 32(5): 70-75.

Eeuwen, M. V. 2017. Mobile conversation commerce: messenger chatbots as the next interface between businesses and consumers. Master's Thesis, University of Twente, Netherlands.

Ekman, P., Friesen, W. & Ellsworth, P. 1972. *Emotion in the Human Face*. New York: Pergamon Press.

FastAPI. n.d. FastAPI. https://fastapi.tiangolo.com/ [3 July 2025].

Fauzi, A., Butar, J. B., Budi, I., Ramadiah, A., Putra, P. K. & Santoso, A. B. 2024. Supervised machine learning entity sentiment analysis: Prediction of support for 2024 Indonesian presidential candidates. *International Information and Engineering Association* 38(2): 587-594.

Fidel, R. 2012. *Human Information Interaction: An Ecological Approach to Information Behavior*. Cambridge: MIT Press.

Fink-Hafner, D., Dagen, T., Doušak, M., Novak, M. & Hafner-Fink, M. 2019. Delphi method: Strengths and weaknesses. *Advances in Methodology and Statistics/Metodoloski Zvezki* 16(2): 1-19.

Fisher, K. E., Erdelez, S. & McKechnie, L. (eds.). 2005. *Theories of Information Behavior*. New York: Information Today, Inc.

Følstad, A. & Brandtzaeg, P. B. 2017. Chatbots and the new world of HCI. *Interactions* 24(4): 38-42.

Fortune. 2017. Those annoying chatbots will save business billions, researcher says. http://fortune.com/2017/05/10/chatbots-business-billions-savings/ [10 December 2017].

Gamage, G., Mills, N., Rathnayaka, P., Jennings, A. & Alahakoon, D. 2022. Cooee: An artificial intelligence chatbot for complex energy environments. *Proceedings of the 2022 15th International Conference on Human System Interaction (HSI),* pp. 1–5.

Gastezzi, C. E. B., Rodríguez, M. M. F. & Castillo, A. 2024. Theoretical foundations on customer experience (customer experience, NPS, CSAT, CES, Service Balcony, Journey Map). *Journal of Business and Entrepreneurial Studies* 8(2): 10-21.

Gawande, V., Al Badi, H., Al Makharoumi, K. & Cain, R. 2021. Study design and implementation of NLP techniques for automated grading of answer: A conceptual model. *International Journal of Innovation in Computational Science and Engineering* 2(1): 1-08.

GeeksforGeeks. 2021. *Introduction to Stemming*. 23 July 2021. from https://www.geeksforgeeks.org/introduction-to-stemming/ [27 September 2022].

Geetha, V., Gomathy, C. K., Narayana, R. V. & Krishna, T. V. 2022. Artificial intelligence chatbot using phyton. *Journal of Engineering, Computing and Architecture* 12(3): 251-259.

Goktas, P. & Agildere, A. M. 2024. Transforming radiology with artificial intelligence visual chatbot: a balanced perspective. *Journal of the American College of Radiology* 21(2): 224-225.

Goli, M., Sahu, A. K., Bag, S., & Dhamija, P. 2023. User's acceptance of artificial intelligence-based chatbots: An empirical study. *International Journal of Technology and Human Interaction (IJTH)* 19(1): 1-18.

Griol, D. & Callejas, Z. 2013. An architecture to develop multimodal educative applications with chatbots. *International Journal of Advanced Robotic Systems* 10(3): 175.

Gu, Q., Wang, Z., Zhang, H., Sui, S. & Wang, R. 2024. Aspect-Level sentiment analysis based on Syntax-Aware and Graph convolutional networks. *Applied Sciences* 14(2): 729.

Gupta, S. n.d. Pre-processing of text data in machine learning. https://www.enjoyalgorithms.com/blog/text-data-pre-processing-techniques-in-ml [4 January 2025].

Hadni, M., Lachkar, A. & Ouatik, S. A. 2012. A new and efficient stemming technique for Arabic Text Categorization. *Proceedings of the 2012 International Conference on Multimedia Computing and System*, pp. 791-796.

Halimah Badioze Zaman. 2018. *Simbiosis Seni, Sains, dan Teknologi Berasingan ke Multimedia-Fusion: Syarahan Perdana*. Bangi: Universiti Kebangsaan Malaysia Press.

Harshith. H. 2024. Text preprocessing in natural language processing – towards data science. Medium, 27 July 2024. https://towardsdatascience.com/text-preprocessing-in-natural-language-processing-using-python-6113ff5decd8 [15 September 2023].

Hassan, A., Sabo, M., Isah, A. A., Auwalu, S., Marzuq, M., Ahmad A. A. & Iliyasu, R. 2021. Sentiment analysis technique for text analytics in scientific literatures domain: A review. *Proceedings of the 1st National Conference on Functional Education*, pp. 1-5.

Hertzum, M., Anderson, H. H., Andersen, V. & Hansen, C. B. 2002. Trust in information sources: Seeking information from people, documents, and virtual agents. *Interacting with Computers* 14(5): 575-599.

Hill, J., Ford. W. R. & Farreras, I. G. 2015. Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. *Computer. Hum. Behav.* 49: 245-250.

Hippner, H. & Rentzmann, R. 2006. *Text Mining*. Ingolstadt: Springer-Verlag.

Hsu, C. C. & Sandford, B. A. 2007. The Delphi technique: Making sense of consensus. *Practical Assessment, Research and Evaluation* 12(10): 1-8.

Hu, M. & Liu, B. 2004. Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 168-177.

Hussein, D. M. E. D. M. 2018. A survey of sentiment analysis challenges. *Journal of King Saud University – Engineering Sciences* 30(4): 330-338.

Ibrahim, A. F. 2018. Internship Application Systems (IAS) for university students using laravel. *Journal of Computing Research and Innovation* 3(4): 12-18.

Iqbal, S. & Pippon-Young, L. 2009. The Delphi method. *Nursing Research* 46(2): 116-118.

Islam, T., Sheakh, M. A., Sadik, M. R., Tahosin, M. S., Foysal, M. M. R., Ferdush, J., & Begum, M. 2024. Lexicon and deep learning-based approaches in sentiment analysis on short texts. *Journal of Computer and Communications* 12(1): 11-34.

ISO/IEC. 2001. ISO/IEC 9126. Software Engineering – Product quality. ISO/IEC.

Jeyapriya, A., & Selvi, C. K. 2015, February. Extracting aspects and mining opinions in product reviews using supervised learning algorithm. *Proceedings of the 2015 2nd International Conference on Electronics and Communication Systems (ICECS)*, pp. 548-552.

Jodha, R. & Dadheech, A. 2019. Analysis and evaluation of unstructured data based on stemming algorithm. *American International Journal of Research in Formal, Applied and Natural Sciences AIJRFAN* 1: 19-201.

Joshi, M., Prajapati, P., Shekh, A. & Vala, V. 2017. A survey on sentiment analysis. *International Journal of Computer Application* 163(6): 34-38.

Kaleem, M., Alobadi, O., O'Shea, J. & Crockett, K. 2016. Framework for the formulation of metrics for conversational agent evaluation. RE-WOCHAT: Workshop on Collecting and Generating Resources for Chatbots and Conversational Agents-Development and Evaluation Workshop Programme, 28th May 2016.

Kavaz, E., Puig, A. & Rodriguez, I. 2023. Chatbot-based natural language interfaces for data visualization: A scoping review. *Applied Sciences* 13(12): 7025.

Kavitha, C. & Kavitha, K. P. 2023. A chatbot system for education NLP using Deep Learning. *Proceedings of the 2023 Eight International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, pp. 1-7.

Khan, M. A. & Alshahrani, M. 2020. Energy-efficient cloud computing: A review. *IEEE Access* 8: 123456-123478.

Kim, Y., Levy, J. & Liu, Y. 2020. Speech sentiment and customer satisfaction estimation in socialbot conversation. *Proceedings of the INTERSPEECH* 2020, pp. 1833-1837.

Knowles, G. & Don, Z. M. 2004. The notion of a "lemma": Headwords, roots, and lexical sets. *International Journal of Corpus Linguistics* 9(1): 69-81.

Köhler, J., Philippi, S., Specht, M. & Rüegg, A. 2006. Ontology based text indexing and querying for the semantic web. *Knowledge-based System* 19(8): 744-754.

Kowcika, A., Gupta, A., Sondhi, K., Shivhre, N. & Kumar, R. 2013. Sentiment analysis for social media. *International Journal of Advanced Research in Computer Science and Software Engineering* 3(7): 216-221.

Krikelas, J. 1983. Information-seeking behavior: Patterns and concepts. *Drexel Library Quarterly* 19: 5-20.

Kuligowska, K. 2015. Commercial chatbot: Performance evaluation, usability metrics and quality standards of embodied conversational agents. *Professionals Center for Business Research* 2(2): 1-16.

Laravel. 2023. Installation - #Meet Laravel. https://laravel.com/docs/9.x#why-laravel [5 May 2025].

Lase, M. 2020. Banking on the future: The role of chatbots in the financial sector. *International Journal of Financial Innovation* 5(3): 210-221.

Le, H. S. & Lee, J. H. 2016. Applying opinion mining for exploring foreign visitors' preferences on hotel services. *Proceedings of the Information Systems in Business and Management Conference (ISBM16)*, pp. 1-8.

Lehvä, J., Mäkitalo, N. & Mikkonen, T. 2017. Case study: Building a serverless messenger chatbot. *Proceedings of the International Conference on Web Engineering*, pp. 75-86.

Lennon, M., Peirce, D. S., Tarry, B. D. & Willett, P. 1981. An evaluation of some conflation algorithms for information retrieval. *Journal of Information Science* 3(4): 177-183.

Lexicon. 2019. Oxford University Press's dictionary. https://en.oxforddictionaries.com/definition/lexicon [16 May 2019].

Li, H. 2017. Deep learning for natural language processing: Advantages and challenges. *National Science Review* 5(1): 24-26.

Liu, B. 2012. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies* 5(1): *1*-167.

Liu, C., Zhang, L., Jiang, Q., Yang, F. & Zhao, Y. 2014. The influence of task type on search experience from the perspective of time constraint. *Proceedings of the American Society for Information Science and Technology*, 5(1), pp. 1-4.

Lu, B., Ott, M., Cardie, C. & Tsou, B. K. 2011, December. Multi-aspect sentiment analysis with topic models. *Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops*, pp. 81-88.

Machová, K., Szabóova, M., Paralič, J. & Mičko, J. 2023. Detection of emotion by text analysis using machine learning. *Frontiers in Psychology* 14: 1190326.

MacKenzie, I. S. 2024. Human-computer interaction: An empirical research perspective. New York: Elsevier Science.

Manju, N., Harish, B. S. & Prajwal, V., 2019. Ensemble feature selection and classification of internet traffic using XGBoost classifier. *International Journal of Computer Network and Information Security* 10(7): 37.

Marketsandmarkets. 2017. Chatbot market worth 3,172.0 million USD by 2021. https://www.marketsandmarkets.com/PressRelease/smart-advisor.asp [10 December 2017].

Marques, M. 2018. Top 3 chatbots that are changing the travel industry. Part 1: The Revolution. 14 March 2018. https://medium.com/hijiffy/top-3-chatbots-that-are-changing-the-travel-industry-d325082c50b8 [4 April 2025].

Maruti TechLab. 2017. How to develop a chatbot from scratch. 9 May 2017. https://chatbotsmagazine.com/how-to-develop-a-chatbot-from-scratch-62bed1adab8c [27 March 2025].

Mateen, A., Yasir, M., Qamar Nawaz, S. A., Yasin, Q. & Yunusi, M. 2021. An analysis on text mining techniques for Smart Literature Review. *International Journal of Advanced Trends in Computer Science and Engineering* 10(2): 1284-1288.

Matharaarachchi, A., Mendis, W., Randunu, K., De Silva, D., Gamage, G., Moraliyage, H., Mills, N. & Jennings, A. 2024. Optimizing generative AI chatbots for Net-Zero emissions energy Internet-of-Things infrastructure. *Energies* 17(8): 1935.

Mccarthy, J. 2007. What is artificial intelligence. Technical report, Standford University, United States.

Mclaren, I. 2022. Vibe Check: An Introduction to Sentiment Analysis. *Summa Linguae*. https://summalinguae.com/data/introduction-to-sentiment-analysis/ [3 July 2025].

Meerschman, H. & Verkeyn, J. 2019. Towards a better understanding of service quality attributes of a chatbot. Master's Dissertation, Ghent University, Belgium.

Merriam-Webster. n.d. Emotion. In. *Merriam-Webster.com/dictionary*. https://www.merriam-webster.com/dictionary/emotion [3 November 2022].

Midda, R.K. 2023. *Maximizing Productivity and Efficiency: Harnessing the Power of AI Chatbots (ChatGPT, Microsoft Bing, and Google Bard): Unleashing Your Productivity Potential: An AI Chatbot Guide for Kids to Adults*. New York: Independently Published.

Mimoun, M. S. B., Poncin, I. & Garnier, M. 2012. Case study – Embodied virtual agents: An analysis on reasons for failure. *Journal of Retailing and Consumer Services* 19(6): 605-612.

Mindbowser. 2017. Chatbot market survey-2017: Mindbowser info solutions 2017. http://mindbowser.com/chatbot-market-survey-2017 [10 December 2017].

Miner, G., Elder, I. V., Fasr, A., Hill, T., Nisbet, R. & Delen, D. 2012. *Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications.* New York: Academic Press.

Miranda, S. M. & Saunders, C. S. 2003. The social construction of meaning: An alternative perspective on information sharing. *Information System Research* 14(1): 87-106.

Moffitt, M. 2018. Chatbots in banking: Improving customer experience. *Journal of Financial Services Marketing* 23(2): 127-136.

Mohammad, S. M. 2017. Challenges in sentiment analysis. In Cambria, E., Das, D., Bandyopadhyay, S. & Feraco, A. (eds.). *A Practical Guide to Sentiment Analysis*, pp. 61-63. Cham: Springer.

Molina, J. 2022. Sentiment Analysis Explained. Symbl.ai, 31 August 2022. https://symbl.ai/blog/sentiment-analysis/ [28 September 2023].

Morente-Molinera, J. A., Kou, G., Peng, Y., Torres-Albero, C. & Herrea-Viedma, E. 2018. Analysing discussion in social networks using group decision making methods and sentiment analysis. *Information Sciences* 447: 157-168.

Morrissey, K. & Kirakowski, J. 2013. 'Realness' in chatbots: establishing quantifiable criteria. *Proceedings of the International Conference on Human-Computer Interaction*, pp. 87-96.

Mujilahwati, S. 2016. Pre-processing text mining pada data Twitter. *Proceedings of the National Seminar on Information and Communication Technology*, pp. 49-56.

Mukherjee, A., Venkataraman, V., Liu, B. & Glance, N. 2013. What yelp fake review might be doing?. *Proceedings of the international AAAI conference on web and social media* 7(1), pp. 409-418.

Neilsen, J. 2006. *Quantitative Studies: How Many Users to Test?*. Berlin: Springer-Verlog.

Newell, A. & Simon, H.A. 1972. *Human Problem Solving*. Upper Saddle River: Prentice Hall.

Nica, I., Tazl, O. A. & Wotawa, F. 2018. Chatbot-based tourist recommendations using model-based reasoning. *Proceedings of the 20th International Configuration Workshop*, pp. 27-28.

Norman, D. A. 2013. *The Design of Everyday Things: Revised and Expanded Edition*. London: Basic Books.

Nuance 2017. Nuance Receives #1 Combined Rating among Enterprise Intelligent Assistance Vendors in Newly-Published Opus Research Report. https://www.nuance.com/about-us/newsroom/press-release/nuance-rated-top-vendor-in-opus-research-report.html [10 December 2017].

Oliver, R. L. 2014. *Satisfaction: A Behavioural Perspective on the Consumer: A Behavioural Perspective on the Consumer*. London: Routledge.

Pandey, M., Williams, R., Jindal, N. & Batra, A., 2019. Sentiment analysis using lexicon based approach. *IITM Journal of Management and IT* 10(1): 68-76.

Pang, B. & Lee, L. 2008. Opinion mining and sentiment analysis. *Foundations and Trends ® in Information Retrieval* 2(1-2): 1-135.

Paramesh, S. P. & Shreedhara, K.S. 2019. Building intelligent service desk system using AI. *International Journal of Science, Technology, Engineering and Management-A VTU Publication* 1(2): 1-8.

Pathak, A., Sharma, S. & Pandey, R., 2021. A methodological survey on sentiment analysis techniques and their applications in opinion mining. *International Journal of Emerging Trends in Engineering and Development* 11(1): 37-45.

Patterson, D., Ganesh, A. & Hsu, W. 2021. Carbon emissions and large neural network training. *Proceedings of the 38th International Conference on machine Learning*, pp. 1-22.

Patterson, D., Gonzalez, J., Holzle, U., Le, Q., Liang, C., Munguia, L., Rothchild, D., So, D. R., Texier, M. & Dean, J. 2022. The carbon footprint of machine learning training will plateau, then shrink. *Computer* 55(7): 18–28.

Pavlíček, A. 2021. Use of chatbots in real life economical. *IT for Practice* 2021: 103.

Peras, D. 2018. Chatbot evaluation metrics: Review paper. *Economics and Social Development: Book of Proceedings*, pp. 89-97.

Poliakova, A. 2010. Application of the Customer Satisfaction Index (CSI) to transport services. *Perner's Contact* 5(4): 208-215.

Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O., et al. 2016. Semeval-2016 task 5: Aspect based sentiment analysis. *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016)*, pp. 19-30.

Porter M. F. 2006. Celebrating 40 years of ICT in libraries, museums, and archives: An algorithm for suffix stripping. *Program: Electronic Library and Information System* 14(3): 211-218.

Pozzi, F. A., Fersini, E., Messina, E. & Liu, B. 2016. *Sentiment Analysis in Social Networks*. New York: Morgan Kaufmann.

Pradhan, V. M., Vala, J. & Balani, P. 2016. A survey on sentiment analysis algorithms for opinion mining. *International Journal of Computer Applications* 133(9): 7-11.

Pressman, R. S. 2010. *Software Engineering: A Practitioner's Approach*. New York: McGraw-Hill International Edition.

Quarteroni, S. & Manandhar, S. 2009. Designing an interactive open-domain question answering system. *Natural Language Engineering* 15(1): 73-95.

Radziwill, N. M. & Benton, M. C. 2017. Evaluating quality of chatbots and intelligent conversational agents. *Journal of Intelligent & Robotic Systems* 91(1): 217–229

Radziwill, N. M. & Benton, M. C. 2017. Evaluating quality of chatbots and intelligent conversational agents. https://www.researchgate.net/publication/316184347_Evaluating_Quality_of_Chatbots_and_Intelligent_Conversational_Agents [3 July 2025].

Rafaeli, S. & Raban, D. R. 2005. Information sharing online: A research challenge. *International Journal of Knowledge and Learning* 1(1-2): 62-79.

Rahman, A. M., Al Mamun, A., & Islam, A. 2017. Programming challenges of chatbot: Current and Future Prospective. *Proceedings of the Humanitarian Technology Conference (RIO-HTC) 2017, IEEE Region* 10, pp. 75-78.

Ramos, R. 2017. Screw the Turing Test – chatbots don't need to act human. 3 February 2017. https://venturebeat.com/ai/screw-the-turing-test-chatbots-dont-need-to-act-human [5 May 2025].

Ranoliya, B. R., Raghuwanshi, N. & Singh, S. 2017. Chatbot for university related FAQs. *Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1525-1530.

Rao, G., Huang, W., Feng, Z. & Cong, Q., 2018. .LSTM with sentence representation for document-level sentiment classification. *Neurocomputing* 308: 49-57.

Razmi, N. A., Zamri, M. Z., Ghazalli, S. S. S. & Seman, N. 2021. Visualizing stemming techniques on online news articles text analytics. *Bulletin of Electrical Engineering and Informatics* 10(1): 365-373.

Rieke, T. D. 2018. The impact of chatbots characteristics on customer satisfaction within the Portuguese Millennial population. Master's Dissertation, Universidade do Porto, Portugal.

Rolnick, D, Donti, P. L. & Kaack, L. H. 2019. Tackling Climate Change with Machine Learning. *ACM SIGKDD Exploration Newsletter* 21(2): 1-4.

Rozite, V., Bertoli, E. & Reidenbach, B. 2023. Data centres & networks - IEA. 11 July 2023. https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks [3 June 2024].

Rusli Ahmad & Hasbee Usop 2011. *Conducting Research in Social Sciences, Humanities, Economics and Management Studies: A Practical Guide.* Kuching: RS Group Publishing House.

Russell-Rose, T. 2017. A framework for chatbot evaluation. https://isquared. wordpress. com/2017/01/24/a-framework-for-chatbot-evaluation [3 June 2024].

Saaty, T. L. 1990. How to make decision: The analytic hierarchy process. *European Journal of Operation Research* 48(1): 9-26.

Saunders, D. 2021. Domain adaptation for neural machine translation. Ph.D's Thesis, University of Cambridge, England.

Sawant, S. 2015. Krikelas' Model of Information Seeking Behavior (1983). *Proceedings of the Information Seeking Behavior and Technology Adoption: Theories and Trends*, pp. 83-93.

Schouten, K. & Frasincar., F. 2015. Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering* 28(3): 813-830.

Selvam, B. & Abirami, S. 2013. A survey on opinion mining framework. *International Journal of Advanced Research in Computer and Communication Engineering* 2(9): 3544-3549.

Semary, N. A., Ahmed, W., Amin, K., Pławiak, P., & Hammad, M. 2024. Enhancing machine learning-based sentiment analysis through feature extraction techniques. *PLoS ONE* 19(2): e0294968.

Serban, I. V., Lowe, R., Henderson, P., Charlin, L. & Pineau, J. 2017. A survey of available corpora for building data-driven dialogue systems. *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2017)*, pp. 3612-3619.

Setiawan, Y., Maulidevi, N. U. & Surendro, K. 2024. The optimization of n-Gram feature extraction based on term occurrence for cyberbullying classification. *Data Science Journal* 23(1): 31.

Shah, K. B., Shetty, M. S., Shah, D. P. & Pamnani, R. 2017. Approaches towards building a banking assistant. *International Journal of Computer Applications* 166(11): 1-6.

Shakirah Mohd Sofi & Ali Selamat. 2022. Implicit aspect extraction techniques in sentiment analysis: A survey. *Malaysian Journal of Information and Communication Technology* 7(1): 18-27.

Shukla, P. K., and Dixit, M. 2015. Big data. In Munir, K., Mubarak, S., Al-Mutairi, and Mohammed, L.A., (eds.). *Handbook of Research on Security Considerations in Cloud Computing*. pp. 326-344. Hershey: IGI Global.

Singh, A. 2025. Energy Consumption and Sustainability of AI-Powered Chatbots / Tools. *Tools* 1: 1-19.

Sirsat, S. R., Chavan, V. & Mahalle, H.S. 2013. Strength and accuracy analysis of affix removal stemming algorithms. *International Journal of Computer Science and Information Technologies* 4(2): 265-269.

Skulmoski, G. J., Hartman, F. T. & Krahn, J. 2007. The Delphi method for graduate research. *Journal of Information Technology Education 6*: 1-21.

Smutny, P. & Schreiberova, P. 2020. Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers and Education* 151: 103862.

Soufyane, A., Abdelhakim, B. A. & Ahmed, M. B. 2021. An intelligence chatbot using NLP and TF-IDF algorithm for text understanding applied to the medical field. *Proceedings of the Emerging Trends in ICT for Sustainable Development*, pp. 3-10.

Su, J., Chen, Q., Wang, Y., Zhang, L., Pan, W. & Li, Z. 2023. Sentence-level sentiment analysis based on supervised gradual machine learning. *Scientific Reports* 13(1): 14500.

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* 9: 502- 510.

Tabassum, A. & Patil, R. R. 2020. A survey on text-processing and feature extraction techniques in natural language processing. *International Research Journal of Engineering and Technology (IRJET)* 7(6): 4864-4867.

Takale, D. G., Patil, A., Jadhav, S., Masram, S., Masarkar, S., Kate, R. & Patil, V. 2024. Sentiment analysis through the application of machine learning algorithms. *Journal of Computer Networks and Virtualization* 2(1): 9-14.

Taylor, R. 1968. Question-negotiation and information seeking in libraries. *College and Research Libraries* 29(3): 178-194.

Text Mining 2020. IBM Cloud Learn Hub. 15 November 2020. https://www.ibm.com/cloud/learn/text-mining [29 September 2022].

Thompson, S. A. 2002. "Object Complements" and conversation towards a realistic account. Studies in Language. *International Journal sponsored by the Foundation "Foundations of Language"* 26(1): 125-163.

Trends, M. 2020. Types of sentiment analysis and how brands perform them. Analytic Insights, 12 November 2020. https://www.analyticsinsight.net/types-of-sentiment-analysis-and-how-brands-perform-them/ [4 September 2023].

Trialopa, E. F. 2022. Optimizing the customer service performance with using chatbot by utilizing bot accuracy. Journey Completion Rate, and Customer Satisfaction Score (CSAT) (Case Study: PT YAI). Master's Thesis. Universitas Islam Indonesia, Yogyakarta, Indonesia.

Tsakiris, G, Papadopolous, C., Patrikalos, G., Kollias, K.F., Asmopolous, N. & Fragulis, G. F. 2022. The development of a chatbot using Convolutional Neural Networks. *SHS Web of Conferences* 139: 03009.

Tumasjan, A., Sprenger, T., Sandner, P. & Welpe, I. 2010, May. Predicting elections with twitter. What 140 characters reveal about political sentiment. *Proceedings of the International AAAI Conference on Web and Social Media* 4(1), pp, 178-185.

Turney, P. D. 2002, July. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pp. 417-424.

Tutorialspoint. 2020. Learn natural language toolkit. https://www.tutorialspoint.com/natural_language_toolkit/natural_language_toolkit_stemming_lemmatization.htm# [3 July 2025].

US Department of Justice. 2003. *The Department of Justice Systems Development Life Cycle Guidance Document*. USA: US Department of Justice.

Valtolina, S. & Neri, L. 2021. Visual design of dialogue flows for conversational interfaces. In *Behaviour and Information Technology* 40(10): 1008-1023.

Van der Heijden, H. 2020. Cost reduction and efficiency through automation in banking. *Journal of Banking and Finance Technology* 4(2): 221-234.

Verstegen, C. 2022. Breaking Down the Pros and Cons of Chatbots as a Customer Experience Solution. Chatdesk, 12 August 2022. https://www.chatdesk.com/blog/ [25 September 2022].

Vinodhini, G. & Chandrasekaran, R. M. 2012. Sentiment analysis and opinion mining: A survey. *International Journal* 2(6): 282-292.

Vogel, J. 2017. *Chatbots: Development and Applications*. Berlin: University of Applied Sciences.

Vohra, S. M. & Teraiya, J. B. 2013. A comparative study of sentiment analysis techniques. *Journal JIKRCE* 2(2): 313-317.

Vu. L, & Le, PhD, T. 2017. A lexicon-based method for Sentiment Analysis using social network data. *Proceedings of the International Conference on Information and Knowledge Engineering (IKE),* pp. 10-16.

Wallace, R. 2003. *The Elements of AIML Style*. New York: Alice AI Foundation.

Wang, W. T., Ou, W. M. & Chen, W. Y. 2019. The impact of inertia and user satisfaction on the continuance intentions to use mobile communication applications: A mobile service quality perspective. *International Journal of Information Management* 44: 178-193.

Wankhade, M., Rao, A. C. S. & Kulkarni, C. 2022. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review* 55(7): 5731-5780.

Weizenbaum, J. 1966. ELIZA – a computer program for the study of natural language communication between man and machine. *Communication of the ACM* 9(1): 36-45.

Whitby, B. 1996. The turing test – AI's biggest blind alley?. *Artificial Intelligence: Critical Concepts* 4(1): 195-203.

Willett, P. 2006. The Porter stemming algorithm: Then and now. *Program: Electronic Library and Information System* 40(3): 219-223.

Wilson, T. D. 1981. On user studies and information needs. *Journal of Documentation,* 37(1): 3-15.

Wilson, T. D. 1997. Information behavior: An interdisciplinary perspective. *Information Processing and Management* 33(4): 551-572.

Wilson, T. D. 2005. Evolution in information behavior modelling: Wilson's model. https://proftomwilson.wordpress.com/wpcontent/uploads/2019/03/2005siguse.pdf [2 February 2025].

Wilson, T., Wiebe, J. & Hoffman, P. 2005. Recognizing contextual polarity in phrase level sentiment analysis. *Proceedings of the EMNLP 2005* pp. 347-354.

Wizu. 2018. A visual history of chatbots. 1 March 2018. https://chatbotmagazine.com/a-visual-hitory-of-chatbots-8bf3b31dbfb2 [24 April 2025].

Xiang, Y., Zhang, Y., Zhou, X., Wang, X. & Qin, Y. 2014, October. Problematic situation analysis and automatic recognition for Chinese online conversational system. *Proceedings of the Third CIPS-SIGHAN Joint Conference on Chinese Language Processing*, pp. 43-51.

Xiao, Y., Li, C., Thürer, M., Liu, Y. & Qu, T. 2022. Towards Lean Automation: Fine-Grained sentiment analysis for customer value identification. *Computers and Industrial Engineering* 169: 108186.

Xu, A., Liu, Z., Guo, Y., Sinha, V., Akkiraju, R. 2017. A new chatbot for customer service on social media. *Proceeding of the ACM Conference on Human Factors in Computing Systems*, pp 1-5.

Xu, J., He, R., Zhou, W. & Li, T. 2018. Trouble ticket routing models and their applications. *IEEE Transactions on Network and Service Management* 15(2): 530-543.

Yaakub, M. R., Latiffi, M. I. A. & Zaabar, L. S. 2019. A review on sentiment analysis techniques and applications. *IOP Conference Series: Materials Science and Engineering* 551(1): 012070.

Yan, M. Castro, P., Cheng, P, & Ishakian, V. 2016. Building a chatbot with serverless computing. *Proceedings of the 1st International Workshops on Mashups of Things and APIs*, pp. 1-5.

Yang, B. & Cardie, C. 2014. Context-aware learning for sentence-level sentiment analysis with posterior regularization. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 325-335.

Yun, Y., Ma, D. & Yang, M. 2021. Human-computer interaction-based decision support system with applications in data mining. *Future Generation Computer Systems* 114: 285-289.

Zafarani, R. & Liu, H. 2014. Behavior analysis in social media. *IEEE Intell. Sys* 29(4): 9-11.

Zahoor, S. & Rohilla, R. 2020. Twitter sentiment analysis using lexical or rule based approach: A case study. *Proceedings of the 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, pp. 537-542.

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

APPENDIX A

PRELIMINARY STAGE QUESTIONNAIRE

PRELIMINARY UNDERSTANDING ON USER TOWARDS CHATBOT

UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

Survey on user understanding and satisfaction level toward existing chatbot.

Dear Participant,

My name is Nurul Muizzah binti Johari, a PhD student from The National University of Malaysia (UKM). I am conducting a research study on improvised chatbot system in tourism sector and investigate the semantic analysis as well as customer satisfaction index (CSI) after using chatbot. I am inviting you to participate in this research by completing the following survey. You have approximately one (1) week to complete and return the survey. If you choose to participate, kindly answer all the questions stated in this survey and email the completed form to muizjohari91@gmail.com

The survey consists of three (3) sections related to demographic information, chatbot evaluation and customer satisfaction index (CSI).

SECTION A: DEMOGRAPHIC INFORMATION
SECTION B: CHATBOT EVALUATION
SECTION C: CUSTOMER SATISFACTION INDEX (CSI)

I appreciate your cooperation and time throughout the survey duration. Your contribution will provide useful input to my research study. Completing and returning this survey form will indicate your willingness to participate in this research. For any further information, you may reach me at 6017-5637551 or muizjohari91@gmail.com

This questionnaire is divided into three sections: SECTION A (Demographic Information), SECTION B (Chatbot Evaluation), and SECTION C (Customer Satisfaction Index-CSI). Participants were required to fill in all the information in Section A. For Section B and Section C, there will be multiple-option statements.

SECTION A: DEMOGRAPHIC INFORMATION

Please tick (/) on the suitable answer below. / *Sila tandakan (/) pada pilihan jawapan yang sesuai di bawah.*

1. Gender / *Jantina:*
   - [ ] Male / *Lelaki*
   - [ ] Female / *Perempuan*

2. Age / *Umur:*
   - [ ] 19- 24 years old / *19-24 tahun*
   - [ ] 25- 30 years old / *25-29 tahun*
   - [ ] 31-34 years old / 30-34 *tahun*
   - [ ] 35 years old and above

3. Occupation
   - [ ] Student
   - [ ] Government sector
   - [ ] Private sector
   - [ ] Self-employed

4. Experience using Chatbot
   - [ ] Siri
   - [ ] Alexa
   - [ ] Cortona
   - [ ] CIMB EVA
   - [ ] Air Asia AVA
   - [ ] Celcom – Emma & Clive
   - [ ] Others

5. Features of Good Chatbot
   - [ ] User Friendly
   - [ ] Answer in brief time
   - [ ] Interaction
   - [ ] Available 24/7
   - [ ] Chatbot personality
   - [ ] Simple interface
   - [ ] Multi-language
   - [ ] Automation on daily life task
   - [ ] Perform simple task without hand

Please carefully read each statement below. Please tick (/) in the space provided according to the given Likert Scale for Section B and Section C.

1	2	3	4	5
Strongly Disagree / *Sangat Tidak Bersetuju*	Disagree / *Tidak Bersetuju*	Neutral / *Neutral*	Agree / *Bersetuju*	Strongly Agree / *Sangat Bersetuju*

SECTION B: CHATBOT EVALUATION

This section consists of statements to evaluate the technical quality of Chatbot. Please tick the answer below according to the Likert Scale above.

No	Statement (s)	1	2	3	4	5
1	The system offers enough interactivity.					
2	The system is easy to use.					
3	It is easy to know what to do each moment.					
4	The amount of information that is displayed on the screen is adequate.					
5	The arrangement of information on the screen is logical.					
6	The chatbot is helpful.					
7	The chatbot is interactive.					
8	The chatbot reacts in a consistent way.					
9	The Chatbot complements the activities without distracting or interfering with them.					
10	The Chatbot provides adequate non-verbal feedback.					

SECTION C: CUSTOMER SATISFACTION INDEX (CSI)

No	Statement (s)	1	2	3	4	5
1	I think that I would like to use this system frequently					
2	I found the system unnecessarily complex					
3	I thought the system was easy to use					
4	I think that I would need the support of technical person to be able to use this system					
5	I found the various function in this system were well integrated					
6	I thought there was too much inconsistency in this system					
7	I would imagine most people would learn to use this system very quickly					
8	I found the system very cumbersome to use					
9	I felt very confident using the system					
10	I needed to learn a lot of things before I could get going with this system					

APPENDIX B

USER ACCEPTANCE TEST

Dear Participant,

We are conducting a research study on an improvised chatbot system and investigating the semantic analysis as well as the customer satisfaction index (CSI). We are inviting you to participate in this research by completing the following test script.

The survey consists of four (4) sections related to demographic information and testing scenarios. chatbot evaluation, and customer satisfaction index (CSI).

SECTION A: DEMOGRAPHIC INFORMATION
SECTION B: TESTING SCENARIO
SECTION C: CHATBOT EVALUATION
SECTION D: CUSTOMER SATISFACTION INDEX (CSI)

I appreciate your cooperation and time throughout the session. Your contribution will provide useful input to our research study. Completing and returning this test script will indicate your willingness to participate in this research.

SECTION A: DEMOGRAPHIC INFORMATION

Please fill out the information below.

1. Gender / *Jantina:*

- [ ] Male / *Lelaki*
- [ ] Female / *Perempuan*

2. Age / *Umur:*

- [ ] 19- 24 years old / *19-24 tahun*
- [ ] 25- 30 years old / *25-29 tahun*
- [ ] 31-34 years old / 30-34 *tahun*
- [ ] 35 years old and above

3. Highest Education */ Peringkat Pengajian Tertinggi*:

- [ ] SPM
- [ ] STPM
- [ ] Diploma
- [ ] Degree
- [ ] Master
- [ ] PhD
- [ ] Other: ______________________________

4. Occupation / *Pekerjaan:*

- [ ] Student
- [ ] Government sector
- [ ] Private sector
- [ ] Self-employed

5. Experience using Chatbot / *Pengalaman menggunakan* Chatbot*:*

- [ ] Yes
- [ ] No

SECTION B: TESTING SCENARIO

Scenario for users

1. Start using the chatbot.
2. Find the type of service provided.
3. Find information related to the loan applications.

Scenario 1	To start a conversation and receive feedback					
Procedure	1. Click "Get Started." 2. Key in any keyword to begin the conversation. 3. Users may use the button for shortcuts. 4. End the conversation.					
Observation		Pass				
		Fail				
Feedback		Very Dissatisfied	Dissatisfied	Neutral	Satisfied	Very Satisfied
	Time taken for chatbot to respond					
	Accuracy of response provided					
Remarks						

Scenario 2	To find the type of services provided by the bank					
Procedure	1. Click "Get Started." 2. Key in any related to banking services. 3. Try to input any unrelated keywords. 4. Observe the response					
Observation		Pass				
		Fail				
Feedback		Very Dissatisfied	Dissatisfied	Neutral	Satisfied	Very Satisfied
	Time taken for chatbot to respond					
	Accuracy of response provided					
Remarks						

Scenario 3	To find information related to the loan application					
Procedure	1. Click "Get Started." 2. Key in any keyword related to loan services. 3. End conversation.					
Observation		Pass				
		Fail				
Feedback		Very Dissatisfied	Dissatisfied	Neutral	Satisfied	Very Satisfied
	Time taken for chatbot to respond					
	Accuracy of response provided					
Remarks						

Please carefully read each statement below. Please tick (/) in the space provided according to the given Likert Scale for Section B and Section C.

1	2	3	4	5
Strongly Disagree / *Sangat Tidak Bersetuju*	Disagree / *Tidak Bersetuju*	Neutral / *Neutral*	Agree / *Bersetuju*	Strongly Agree / *Sangat Bersetuju*

SECTION B: CHATBOT EVALUATION

This section consists of statements to evaluate the Chatbot. Please tick the answer below according to the Likert Scale above.

No	Category	Statement (s)	1	2	3	4	5
1	Functionality	Able to interpret keywords or instructions accurately					
		Flexible to interpreting knowledge					
		Accuracy of the response provided					
		The information provided is secured					
2	Efficiency	Easy to use					
		Quick response					
		Can be used at anytime and anywhere					
		Able to be accessed on any platform					
3	Technical Satisfaction	Able to convey the greeting					
		Able to provide emotional information using tones and expression					
		There is no need for technical skills of proficiency to use the chatbot					
		The chatbot provides information in a fun, interesting, and innovative way					
4	Humanity	Realness of robot					
		Able to create enjoyable interaction					
		Able to maintain the theme of the conversation					
		Can mimic the human personality					
5	Effectiveness	Linguistic accuracy					
		Have a wide range of knowledge (related to banking)					
6	Ethics	Trained with knowledge culture and ethics of users					
		Sensitive to user's concern					

SECTION C: CUSTOMER SATISFACTION INDEX (CSI)

This section consists of statements to evaluate the satisfaction level after using Chatbot. Please tick the answer below according to the Likert Scale above.

No	Statement (s)	1	2	3	4	5
1	I think that I would like to use this system frequently					
2	I found the system unnecessarily complex					
3	I thought the system was easy to use					
4	I think that I would need the support of a technical person to be able to use this system					
5	I found the various function in this system were well integrated					
6	I thought there were too many inconsistencies in this system					
7	I would imagine most people would learn to use this system very quickly					
8	I found the system very cumbersome to use					
9	I felt very confident using the system					
10	I needed to learn a lot of things before I could get going with this system					

APPENDIX C

LIST OF PUBLICATIONS

Journals

1. Johari, N. M., Nohuddin, P. N., Baharin, A. H. A, Yakob, N. A. & Ebadi, M. J. 2022. Feature requirement elicitation process for designing a chatbot application. *IET Networks* 2022: 1-15.
2. Johari, N. M. & Nohuddin, P. N. 2021. Quality attributes for a good chatbot: a literature review. *International Journal of Electrical Engineering and Technology (IJEET)* 12(7): 109-119.
3. Johari, N. M. & Nohuddin, P. N. 2021. Developing a good quality chatbot. *Journal of Information System and Technology Management* 6(22): 88-102.

Proceedings

1. Johari, N. M., Zaman , H. B., Baharin, H., & Nohuddin, P. N. 2023. A visual-based energy efficient chatbot: Relationship between sentiment analysis and customer satisfaction. *International Visual Informatics Conference*, pp. 606-628.
2. Johari, N. M., Zaman, H. B. & Nohuddin, P. N. 2019. Ascertain quality attributes for design and development of new improved chatbots to assess Customer Satisfaction Index (CSI): A preliminary study. In *Advances in Visual Informatics: 6th International Visual Informatics Conference, IVIC 2019*, pp. 135-146.