SSURJET

Sir Syed University Research Journal of Engineering & Technology
2023, Vol. 13, No. 1,
https://doi.org/10.33317/ssurj.508

Comparative Analysis of Classical and Neural Networks based ChatBot's Techniques

Junaid Javed[1], Ahthasham Sajid[1], Imran Ullah Khan[2*], Shahnoor[1], and Iqra Tabassum[1]

[1]*Department of Computer Science, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan*
[2]*Department of Electrical Engineering, Hebei University, Baoding China*

**Correspondence Author: Imran Ullah Khan (khan@hrbeu.edu.cn)*

Received June 11, 2022; Revised January 17, 2023; Accepted January 29, 2023

Abstract

Conversational Agents like Alexa from Amazon, Siri from Apple, Assistant from Google, and Cortana from Microsoft, all demonstrate extraordinary research and potential in the field of conversational agents. A conversational agent, chatter-bot, or ChatBot is a piece of computer software that is supposed to communicate at a level of intelligence comparable to that of a person. ChatBots are designed for a variety of purposes, such as task-oriented helpers and as creators of open-ended discourse. Numerous approaches have been studied on this topic, from primitive types of hard-coded response generators to contemporary ways of constructing artificial intelligence. These are classified as rule-based or neural network-based systems. Unlike the rule-based technique, which is based on pre-defined templates and responses, the neural network approach is based on deep learning models. This article begins with an overview of ChatBots before diving into the specifics of a variety of traditional, rule-based, and neural network-based techniques. A table summarizing previous field research closes the survey. It looks at the most recent and important research on the subject, as well as the evaluation instruments used, areas for improvement, and the applicability of the proposed methods.

Index Terms: *Artificial Intelligence Markup Language, ChatBots, Long Short-Term Memory, Pattern Recognition, Seq2Seq.*

I. INTRODUCTION

While many would believe that the term 'ChatBots', has only gained popularity in recent years, the concept stretches all the way back to the early days of computer interaction. Even before the invention of personal computers, the first ChatBot was introduced. Eliza was founded in 1966 by Joseph Weizenbaum at MIT's Artificial Intelligence Laboratory. Eliza parsed the input for keywords and then triggered the output based on a specified set of rules. Numerous Chatbots are still generating output in this fashion. The history of ChatBot are depicted in figure I [1]. In an attempt to replicate paranoid schizophrenia symptoms, Kenneth Colby created Parry, a Chatbot, using Stanford University physician Kenneth Colby's code [2]. Then there was the A.L.I.C.E., created in 1995 by Richard Wallace. While A.L.I.C.E was a three-time winner of the Loebner Prize, it failed the Turing test. A Turing test determines if a machine can think rationally in the same way as humans do [3]. Subsequently, a slew of virtual assistants was introduced. Siri, developed by Apple, was the first to include conversational assistance. The concept gained popularity, and Google introduced its Google Assistant for Android shortly thereafter. Microsoft quickly followed suit with the creation of Cortana. To take this a step further, smart speakers were designed, enabling human-to-human verbal conversation.

In 1950, Alan Turing responded to the question, "Can machines think?" [4]. Since then, experts in Artificial Intelligence (AI) have been challenged with developing programmes that can simulate human thought in computers. As consultants, talkative companions, or utility software, ChatBots became involved. Numerous design techniques have developed throughout history. This article's objectives are to explain ChatBot's development and provide an example. The main goal of a ChatBot is to produce a suitable response based on natural language input from a human. There are various ways to provide that response, as shown in figure II, which determines how the ChatBots are modeled. The first is a rule-based system that meticulously evaluates user input using hardcoded phrases and pre-made templates to produce the solution. With the rise of deep learning, the alternative approach, based on neural networks, became feasible. The neural network is trained on enormous amounts of data in order to deliver replies that are appropriate for the context and grammatically sound. You can enter data using text, images, or speech. Convolutional Neural Network (CNN) models have been developed for Chatbots to infer pertinent information from photographs, and models for voice-to-text translation [5], and [6].

You can essentially categorize the neural network-based method as either retrieval-based or generative-based. Responses are generated utilizing retrieval-based techniques. Either the most pertinent response is determined using a scoring function, such as conditional probabilities implemented using a neural network, or the context-candidate relationship is examined using reinforced co-ranking [7-9]. The generative approach, in

contrast, generates one word for each input after computing probabilities across the entire vocabulary [10]. A method for combining retrieval and generative models has also been presented [11]. This technique is used to incorporate a reaction into a generative model. The final solution was then obtained by rearranging the retrieved and created responses.

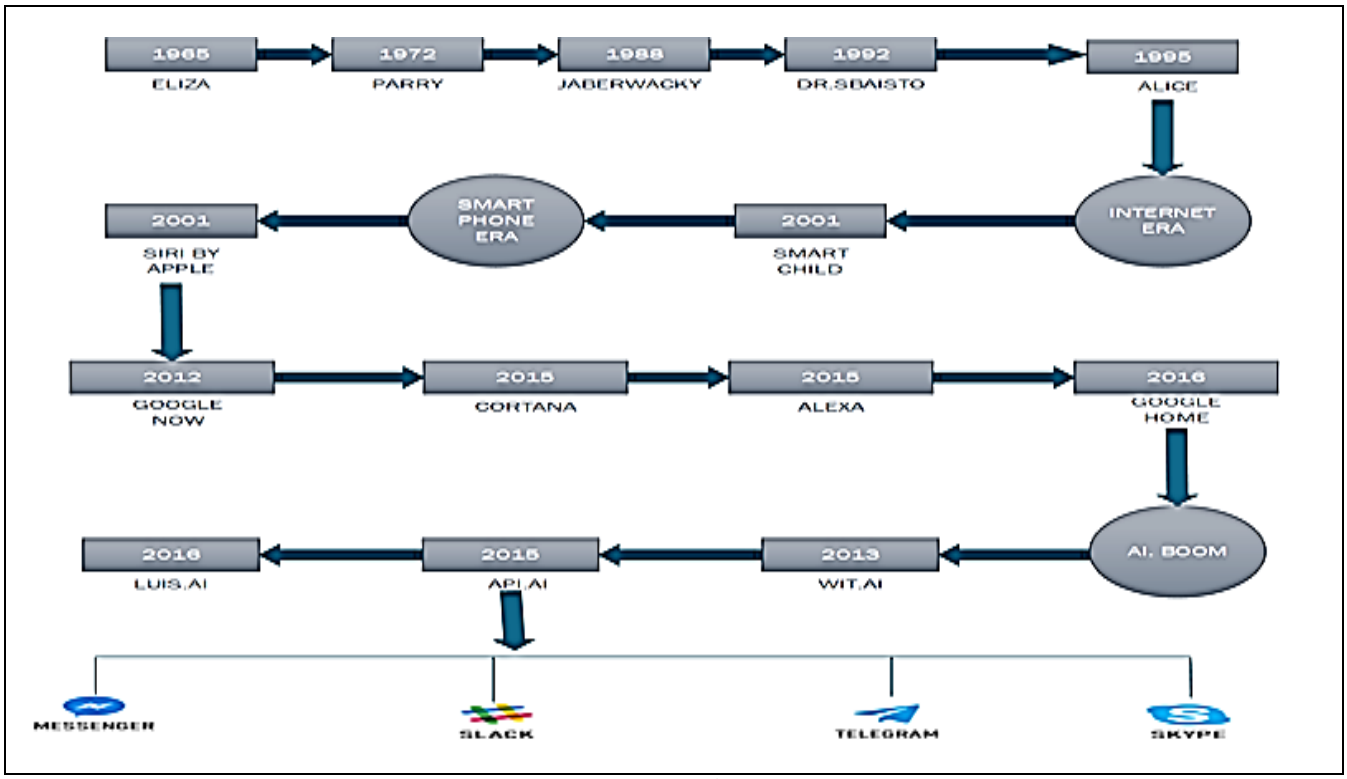

Figure I: History of ChatBots

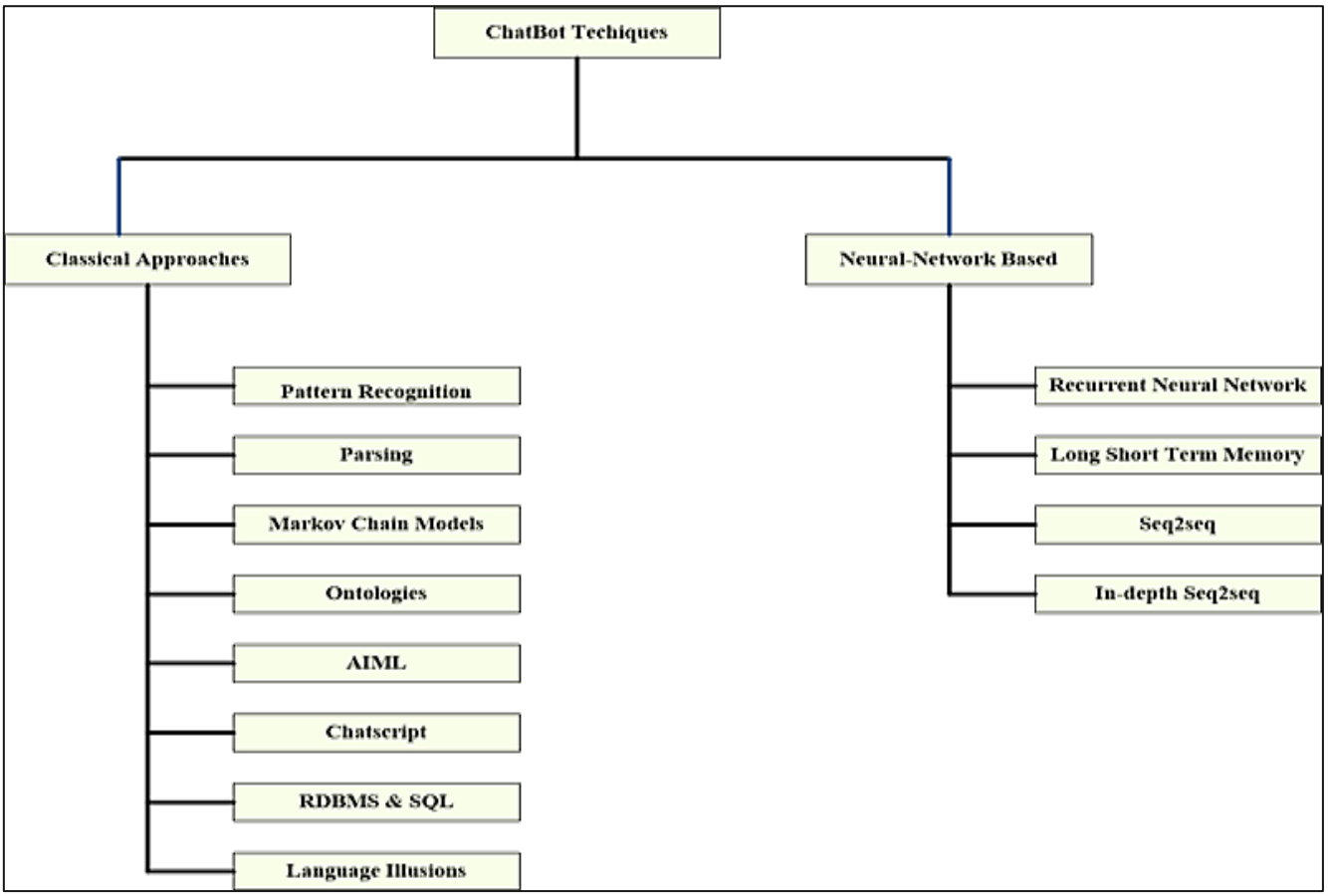

Figure II: Approaches of ChatBot Techniques (see table I, and table II)

ChatBots are classified into two types based on their capabilities; One is goal-oriented, which is not the best characteristic of a conversational AI but is capable of doing specified tasks and responding to domain-specific directions. These ChatBots' primary purpose is to deliver information. There are three types of ChatBots; informative, conversational, and task-oriented as shown in figure III.

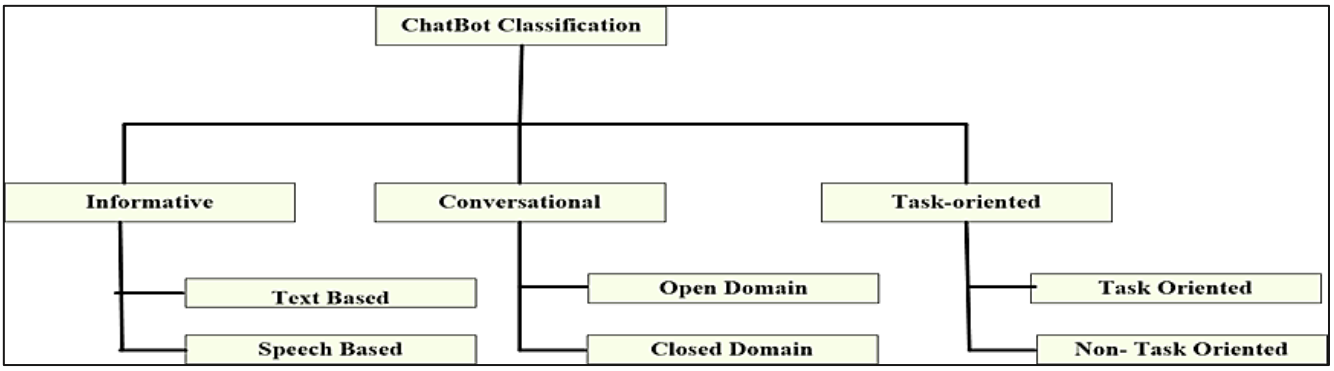

Figure III: Classification of ChatBot

Research Contributions:

- Recognize different techniques of ChatBot.
- Identify the pros and cons of all these techniques.
- Distinguish global metrics parameters of ChatBot based on these techniques.
- Make a comparison analysis according to the related parameters.
- Preprint the future development in the field of ChatBot Techniques

II. RELATED WORK

ChatBots have many real-world applications. E-learning, marketing, medical diagnostics, cultural heritage, e-customer support, and task management all use ChatBots. Internet activity expands its application. E-learning chatBots can help engage students and tailor instruction to their needs. Latent Dirichlet Allocation (LDA) can be used to analyze the user's inquiry, remove stop words, and extract keywords from the conversational transcript to create an e-learning ChatBot [12]. After that, an ontology of learning ideas is generated by course, lesson, subject, or user. Petri net-based ChatBots for type recommendation have been researched [13]. It creates a type of ontology. The Petri net represents the vehicle's genetic makeup based on user inputs of its type (car, scooter, etc.), model number, and production year. After the survey, the weighted total of all responses in each context is calculated. The system avoids repeating inquiries if the weighted total of the contexts is less than the amount of information it already knows. ChatBots also aid in medical diagnosis. Any such communication requires patient clinical information. To stay current on disease diagnosis and treatment, such a system needs ongoing training. Medical technology is essential, and chatBots still need improvement [14]. Based on the tourist profile and context, a ChatBot proposes tourist sites, information, and services. It also recommends local hotels and restaurants, i.e., this ChatBot guides tourists. The architecture revolves around the inference engine. The Context Dimension Tree (CBT) analyses user material and gives suitable responses [15]. Building a ChatBot requires extensive research and learning. ChatBots need context, person, and location information to work efficiently. 'Curious Cat' used crowdsourcing to acquire user data by selecting the most relevant group, assessing dialogue quality, and ensuring consistent replies. This ChatBot also assists couples in private talks [16]. Recently, academics and corporations have advanced ChatBot's use and development. A ChatBot's behavior, appearance, and creator's attributes, as well as privacy and security concerns, affects its trustworthiness. The user's faith in the ChatBot depends on its visual appearance, name, personality, and language skills. Emotion is another crucial part of humanizing a ChatBot, and various attempts have been made to create an emotional ChatBot. The ChatBot's ability to converse as humans improves with AI. Human-to-human ChatBot communication differs qualitatively and substantively from the human dialogue. Human-to-human ChatBot conversations can last a long time. People speak in short, harsh syllables. Empathy distinguishes humans from ChatBots. Humans grasp conversational context better than ChatBots. ChatBots can now understand human emotions. Personal voice assistants embedded into smartphones or specialized home speakers that recognize voice commands, speak in digital voices, and monitor home automated devices, calendars, and emails, i.e., advanced AI chatBots are being created. Siri, IBM Watson, Google Assistant, Cortana, and Alexa are the most popular voice assistants. Several lesser-known voice assistants perform the same basic duties but have different personalities. They have Internet connectivity and produce more results faster than their predecessors. Virtual assistants will soon become essential for everyday life.

III. CLASSICAL APPROACHES TO CHATBOT TECHNIQUES

Rule-based techniques are referred to as such because they create replies based on specified rules, which distinguishes them from traditional procedures. Over time, these requirements have gotten more difficult to meet and more complex to understand. These strategies are most effective when the domain of the conversation is well defined, i.e., when the conversation is focused on a certain topic or activity. In contrast, rule-based systems become inefficient when the input grows more natural or the domain becomes more open, as is the case with Artificial Intelligence (AI).

The AI language, Artificial Intelligence Markup Language (AIML), which is based on XML, allows for the construction of rules. It is the goal of this paper to look at eight common techniques used in classical methods as shown in figure IV.

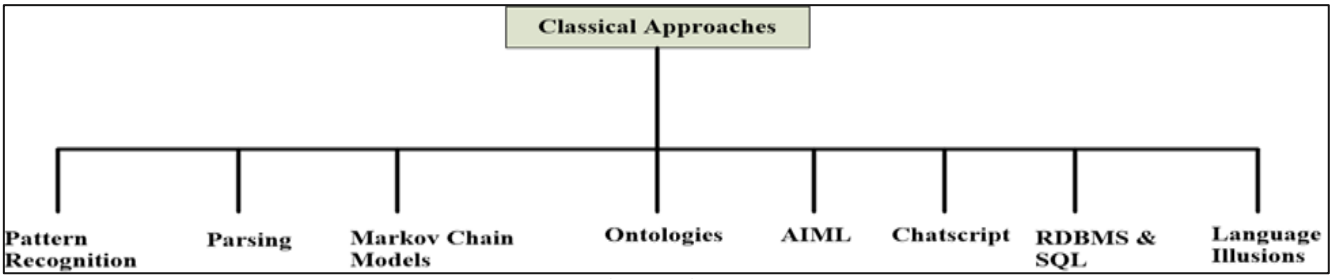

Figure IV: Classification of Classical Approaches

A. Pattern Recognition

When it comes to building ChatBots, pattern recognition is a key technique that is used in practically every ChatBot to some degree. With this method, a response is generated by applying a preset set of rules and templates. ELIZA was the world's first ChatBot created in this manner [17]. In the beginning, it identifies the keywords that are present inside the text by reading it from left to right. Each keyword has a rank or precedence that determines how important it is. Following that, the input string is deconstructed into a template by applying a preset pattern to the string data. For example, when given the input; “I am unhappy”, the system creates a response based on the keyword "sad"—how long have you been "sad”? A large number of people signed up to participate in the ELIZA psychotherapy programmes after it was first introduced in 2012.

B. Parsing

In computer science, parsing is the process of extracting the grammatical structure of a text from it. Initially, the string is separated into bits representing nouns and verb phrases. The nouns, adjectives, and articles are recognized, and a syntax tree is constructed from them. Parsing is a technique that allows the grammatical structure of a phrase to be confirmed in terms of the language in which it was originally written. Previously, basic parsers were employed to detect keywords in text documents. If you want to take food, you can say "take the food" and "can you retrieve the food". Both of these phrases are processed as "take the food". When used with polymorphic input strings, this enables ChatBots with a restricted set of templates and patterns to respond to them. Later ChatBots rely significantly on natural language processing technologies to communicate with users. Parsing may be divided into three categories; syntax, semantics, and pragmatics [18]. Syntax is the first of these categories. Use this strategy in business settings where more control over the flow of speech is needed [19].

C. Markov Chain Models

When expressed mathematically, the ‘Markov Chain Model’ is a model that describes the likelihood of future events given the state of previous affairs. It takes into consideration the possibility that a letter or word may appear in a dataset. It makes use of this probability distribution in order to choose the answer phrases that are most likely to be given. As a result of the sequence in which the Markov chain is constructed, the maximum number of consecutive letters or words that may be entered is determined. When given a string ‘khdddkhhdd’, the probability of meeting the letter ‘k’, is 2/10, the likelihood of meeting the letter ‘h’, is 3/10, and the likelihood of encountering the letter ‘d’, is 5/10, for a zero-order Markov chain. When determining the fixed probability for an order 1 Markov chain, the primary elements will be taken into consideration in addition to the current part. It is supplied with a string that says, "The black dog jumped into the swimming pool". In an order 2 Markov chain, the term, "the black", becomes the word, "dog", the word "black-dog", becomes the word, "jumped", and so on for the remaining words [20], and [21].

D. Ontologies

Ontologies (also known as Semantic Nets) are a sort of organized representation of real-world concepts that are used in computer science and engineering. Classes are also referred to as notions in other instances. They are the subject of the vast majority of ontology definitions. Using ontology, the knowledge base is composed of the cases of various classes linked together to provide an overall picture. For example, a bread class comprises many types of bread available [22]. There are a variety of different features of the bread that may be referenced, including the texture, color, and company. Additional explanations can be provided for the various 'facets' of the slots. These specify, among other things, the kind of value, the cardinality, and the range of the place. Through the use of certain reasoning principles, it is advantageous in that it allows for node searching as well as for the implication of new replies. The ontologies ‘Open Cyc’ [23], and ‘Wordnet’ [24], were employed in the development of ChatBot.

E. Artificial Intelligence Markup Language (AIML)

AIML is a technical invention that aims to make the production of ChatBots more expedient by using machine learning [25]. The stimulus-response paradigm is used to depict a discussion between a ChatBot and a human being. A.I.M.L. is built on the ideas of pattern recognition and matchmaking. Because it is closely connected to XML, it is straightforward to implement, and the tags make communication easier between parties. The Graph master algorithm is in charge of keeping track of the tree that is formed by putting A.I.M.L. patterns in memory. In terms of both space and time efficiency, it is the most effective

solution. Aside from that, it is highly reusable due to its simplicity as well as the availability of both the source code and the associated documentation. A.L.I.C.E. was the world's first ChatBot powered by artificial intelligence, and it still holds the record [26]. A botmaster oversees A.L.I.C.E.'s learning process, which implies A.L.I.C.E. learns while being monitored by a person. Following the initial design of A.L.I.C.E., a large number of other ChatBots were created and developed using the A.I.M.L. programming language.

F. Chat Script

Creating ChatBot is accomplished through the usage of the programming language known as, 'Chat Script'. It was started in 2010 by Bruce Wilcox. In 2010, he received the Loebner Prize for his work on Suzette, a ChatBot that he designed [27]. Chat script is a dialect of the A.I.M.L. programming language. As an alternative to looking through hundreds of categories, chat script searches for a context that is connected to another context. This form of setting is referred to as a "concept", and each concept has its own set of laws that must be followed by everyone. This table is beneficial in reacting to user input by running a basic search within it, which is handy for responding to user input. A food idea is being promoted (including bread and juice as well as vegetables and fruits, as well as pizza, hamburgers, and cold beverages); (I like pizza). Are you a foodie who enjoys trying new things?

G. Relational Database Management System (RDBMS) and Structured Query Language (SQL)

A Relational Database Management System (RDBMS) is used in order to build the ChatBots. The database's main job is to store information about prior discussions and deliver distinct answers to similar questions that have been asked at different times in the past, regardless of when they were asked. SQL is the database management system programming language that is most often used. This method was used in the development of Virtual Diabetes Physician (ViDi) [28].

H. Language Illusions

Other techniques while developing ChatBot, i.e., several popular linguistic tricks include the following:

- Mistypes and Incorrect Keystrokes: When a user enters input, he or she periodically checks the ChatBots to see how it is constructing the answer to the information. Apparently, because they are also human traits, false backspaces, and occasional spelling errors seem to be very common in people, too.
- Pre-Programmed Answers: There are specific patterns that the ChatBot's pattern-matching algorithm is unable to cover. The developer hardcodes such responses.
- Personal History: Developers give the ChatBots, a personality in order to bolster their credibility. It is inculcated with information regarding its birth, age, parents, tastes, and stories [29].

IV. NEURAL-NETWORK-BASED CHATBOT TECHNIQUES

ChatBots powered by neural networks have done away with the time-consuming task of creating rules for each utterance-response combination. There are two methods by which an artificial neural network might deliver a response; it can construct the response from scratch (generative) or it can obtain the response from a large dataset (retrieval or retrieval-based). In addition, various hybrid approaches that blend the best of both worlds have been developed and implemented. A discussion has been started over the underlying structure/model upon which all techniques are built, see figure V.

A. Recurrent Neural Network

When creating a response, any conversational agent would benefit from the capacity to take past conversations and context into consideration, as well as the ability to respond appropriately. When creating a response, it just considers the present input, which results in replies that are repetitive in nature. Because of the Recurrent Neural Network (RNN), it is possible for the ChatBots to use the previously created output as input and provide a more logical answer. RNNs, in contrast to ordinary neural networks, are capable of retaining data over time. Here 'A' is a small component of the neural network seen in figure VI, and 'xt' is its input; it is responsible for the generation of the output 'yt', [30].

B. Long Short Term Memory (LSTM)

A simple R.N.N. has trouble recalling information due to the excessive number of unrolling steps. Weights and activation functions statistically affect gradient value (basically their derivative). The gradient decreases as '1' approaches to '0'. Activation functions like 'tanh', and 'sigmoid', often have derivative values close to zero, making this problematic. In this case, L.S.T.M. fixes fading gradients. The L.S.T.M. activation function is identity, which has a derivative of 1, preventing the back-propagated gradient from reducing. Gates (small code blocks) help the L.S.T.M. The gate controls the data flow. Gates emit a number between 0 and 1 when engaged, indicating how much of each component should pass. This '0' means no data should be sent, while '1' means all data must be sent. The LSTM's input and forget gates control information flow between network units in a dispersed network. These gates define the network state update [31]. The Gated Recurrent Unit (GRU) is popular [32]. A single 'update gate' replaces the input and forget gates in this system.

C. Seq2seq

The 'Seq2seq', a leading machine translation technique, may be useful for conversion models [33]. Figure VII shows that two R.N.N.s form a seq2seq model. R.N.N.s are often made using L.S.T.M. or G.R.U. This function calculates the conditional probability of p (y1, y2,... yn' / 1,2,... xn), where 'x', and 'y' are the input and output sequences, respectively. Function definition; adjust n and n' lengths as needed. The Seq2Seq algorithm works well because two RNNs generate the input and output sequences. Encoder-decoder systems do this.

The first RNN encoder processes the input sequence, generating a vector. This vector initializes R.N.N.2. After that, the decoder's R.N.N. output is given a suitable probability function. Back-propagation adjusts weights while training these two networks in parallel. Seq2seq converts a source language phrase to word embedding vectors. R.N.N.2 translates languages.

This strategy can be modified for chatBots by using the input sentence as the source language string and the answer as the destination language string element.

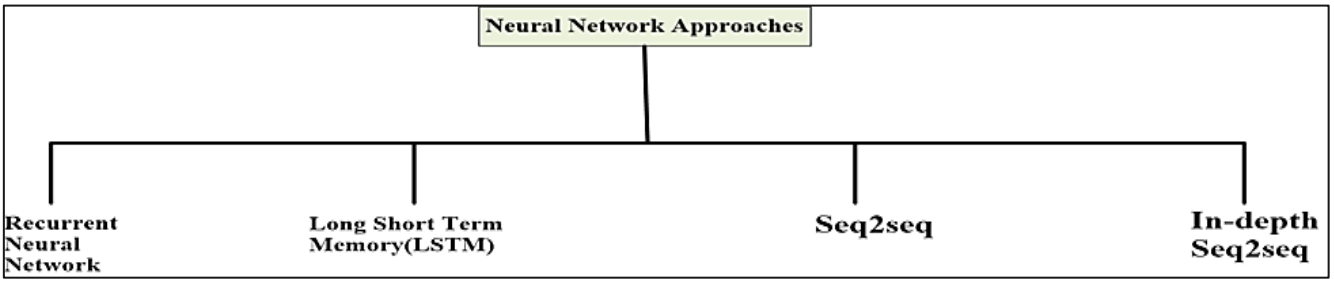

Figure V: Classification of Neural Network Approaches

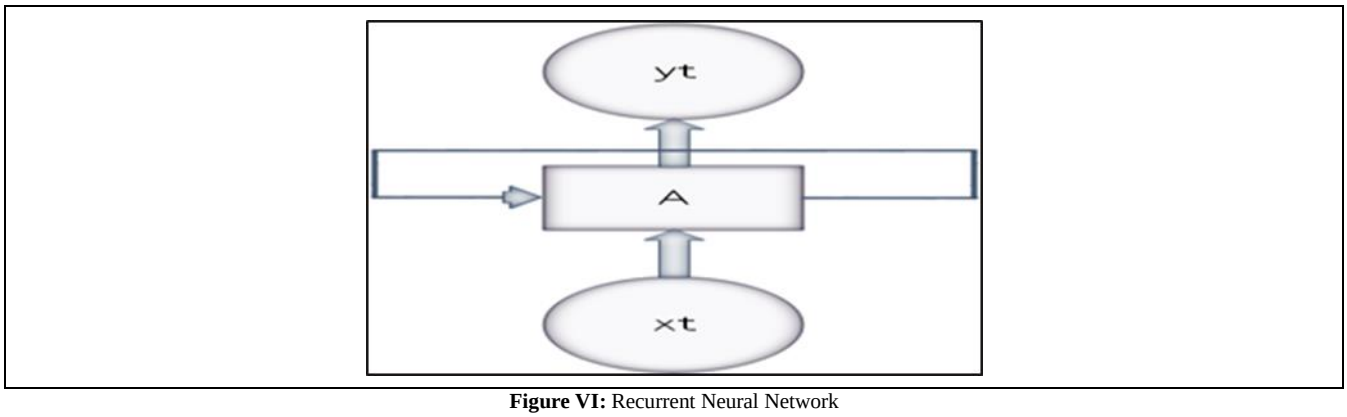

Figure VI: Recurrent Neural Network

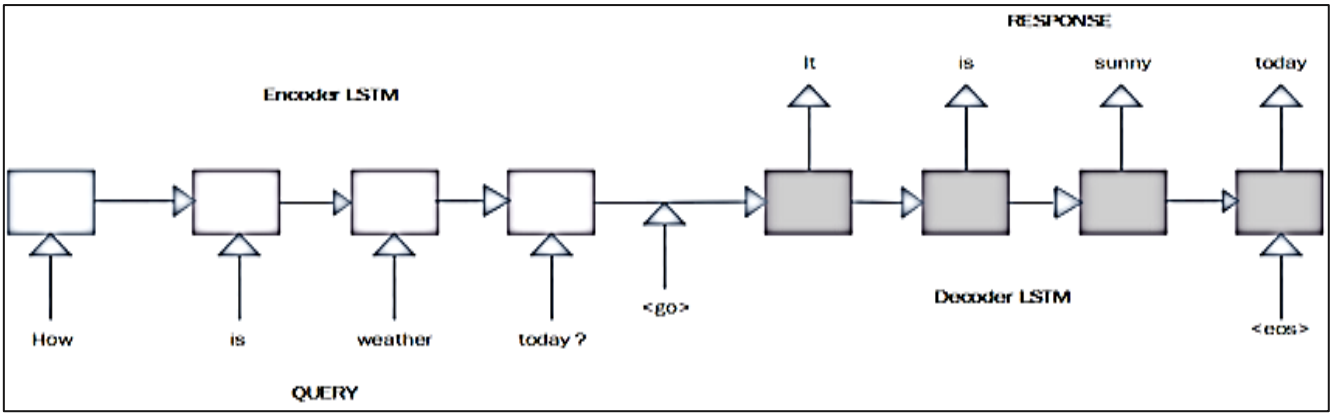

Figure VII: Seq2seq

V. CRITICAL ANALYSIS OF CLASSICAL APPROACHES FOR CHATBOTS

The critical analysis of classical approaches for ChatBots are presented in table I below:

Table I: Classical Approaches of ChatBots and their Analysis

Title & Year	Technique	Pros	Cons	Applications
Communicating with Bots: Symbiotic Agency and the Tay Case (2016) [1]. Dialogues with Colorful Personalities of Early AI (1995) [2]. A.L.I.C.E. Al Foundation (2002) [3].	**Pattern Recognition**	The patterns can be read by humans so that the input modeling phase can be somehow straightforward [1]. Make a significant contribution to the rule	Typically, rule-based systems do not provide new replies, as the developer's knowledge is encoded in the form of conversational patterns [4].	E.L.I.Z.A., PARRY, P.C. Therapist III, HEX, Albert, A.L.I.C.E. and Jabberwacky [6], and [7].

Automatons and artificial intelligence. In: Turing test parsing (2009) [4]. A Chatbots with deep reinforcement learning (2017) [5].		selection process and the answer format [2]. Simple pattern matching method and also simple template pattern for input and output representation [3].	The responses are automatic, repetitive, and lack the uniqueness and spontaneity associated with human responses [4]. There is a rapid response time since no in-depth syntactic or semantic analysis of the input text is undertaken.	
Language development through multi-agent games: learning to interact through symbol sequences (2016) [16]. Human-to-Machine Conversation Modeling with a Chatbots Using Database Knowledge (2016) [35]. A Survey of Techniques for Creating ChatBots in Speech Conversation Systems (2015) [36]. An examination of the ChatBot system via a Loebner prize competition (2012) [37].	**Parsing**	Parsing assists in validating the grammatical structure of a sentence in relation to a language [35]. This strategy is employed in corporate situations that need greater control over the conversational flow [16]. This is an important step toward making sense of meaningful sentences [36].	Ambiguity happens when a sentence's language or grammar allows for various readings [37]. When ambiguity exists, the probability decreases and a considerably narrower range of probabilities may exist [38]. Misleading parsing can lead to wrong results [39].	Jabberwacky [16] Cleverbot, Chatfuel, and Watson [40].
E.L.I.Z.A.—a computer application for the study of human-machine natural language communication (1966) [17]. Natural language processing tutorial (2007) [18]. A.l.anatomy I.C.e's (2009) [41].	**Markov Chain Models**	It utilizes this probability distribution to determine the most likely response phrases. Because Markov models are entirely generic, the produced sequences appear to represent a sample of real-world usage, as long as the model accurately describes operational behavior [41]. The model is based on a formal stochastic process that has a well-defined analytical theory [41].	The number of states in a complicated system tends to increase exponentially [41]. As a fallback mechanism, it used to make a nonsense statement that used to sound natural [17].	HeX [17] MegaHAL[18].
An examination of Chabot systems via the lens of a Loebner prize competition (2012) [20]. MegaHAL II I! II. Computer Introduction (2000) [21]. Using Deep Neural Networks to Learn to Respond to a Retrieval-Based Human-Computer Conversation System (2016) [43].	**Ontologies**	When instances of multiple classes are integrated with the ontology, the knowledge base is formed [19]. In addition to searching through the nodes, specific reasoning rules might infer additional replies [19]. Provide unique scalability and interoperability properties [42].	Ontology-based ChatBots and their inability to respond to situations outside of their knowledge database [43]. The ontology starts to become large, and finding a path between related nodes begins to take asymptotically more time [44].	OpenCyc [20] Wordnet [21].
A.L.I.C.E.'s anatomy (2009) [26] A quick tutorial on the artificial intelligence markup language (2013) [39]. A review of the open-source ChatBots languages A.I.M.L. and chat script (2017) [40]. A.l.anatomy I.C.e's (2009) [41]. An Interactive ChatBot Powered by A.I.M.L. (2020) [42].	**AIML**	Contribute to the development of the Knowledge Base for ChatBots that use the Pattern Matching technique [39]. Enhances usability, facilitates learning and execution, and makes pre-authored A.I.M.L. collections available [40]. A.I.M.L. matches the best pattern using a pattern-matching approach that performs a first-depth search in the Graph master [41].	The knowledge is provided in the form of an A.I.M.L. file collection. If knowledge is created using data acquired from the Internet, it will not be automatically updated and will need to be updated on a frequent basis [42]. A.I.M.L. is famously difficult to maintain due to its restricted collection of matching patterns [44]. Creating a suitable knowledge base may be time-consuming, costly, and exhausting [43].	A.L.I.C.E. [26] PABX (Private automated branch exchange) [44] Humorist bot [45].

Statistical machine translation: Learning phrase representations using an R.N.N. encoder-decoder (2014) [8]. Bringing it to life: The Loebner-award-winning ChatBots design (2014) [46]. An overview of conversational agent design approaches (2017) [47].	**Chat script**	An embedded tagger and parser examine user input and makes it more grammatically, syntactically, and semantically correct [46]. Developers may easily use pre-built databases of concepts, which simplifies the process of developing ChatBots [47]. ChatScript also offers long-term memory in the form of variables that may be used directly or in conjunction with conditionals to generate Chatbot answers [47].	The scripting language used by ChatScript is more sophisticated than A.I.M.L.'s parsed line-delimited languages [8]. ChatScript is capable of combining many rules in complex ways and providing replies that cannot be described in A.I.M.L [8]. RiveScript code can be converted to ChatScript but not vice versa [8].	Suzette, Rosette, Mistsuku, and Chip Vivant [48].
A significant investment in the infrastructure of knowledge (1995) [23]. A.L.I.C.E.'s anatomy (2009) [26]. Techniques for designing ChatBots are discussed. Journal of International Computer Applications (2018) [49]. Google is now geared toward the extroverted, whereas Cortana is geared toward the introverted (2017) [50].	**RDBMS & SQL**	The primary reason for utilizing the database is to recall earlier talks and provide unique responses to identical queries given at various points in time. The database was used to store knowledge, and the interpreter was used to store function and procedure sets for the pattern-matching requirement [49]. This strategy ensures the dialogue's continuity and accuracy by allowing the dialogue system to access some past information history [50].	As the amount of data in a relational database grows, the system becomes increasingly difficult [50]. This information can be utilized to move data across systems. This might result in data loss [50].	ViDi (Virtual Diabetes physician) [26] MultiWOZ [23].
Chabot Modeling Human-to-Machine Conversations Using Database Knowledge (2016) [35]. A Survey of Techniques for Creating ChatBots in Speech Conversation Systems (2015) [36]. Thai-FAQ ChatBots powered by RNN-LSTM (2018) [51].	**Language Illusions**	Increase the variety and persuasiveness of the knowledge base [36]. Language tricks are employed to accomplish a certain objective and to give other responses to queries [51].	Might struggle with understanding customer requests and provide incorrect or insufficient answers [52]. These tricks simulate stereotyped responses, typing errors, and even irrational responses [35].	Google Now Cortana Google Home [36].

VI. CRITICAL ANALYSIS OF NEURAL NETWORK-BASED APPROACHES

The critical analysis of neural-network-based approaches of ChatBots are presented in table II below:

Table II: Neural-Network-Based Approaches and their Analysis

Title & Year	Technique	Pros	Cons	Applications
The vanishing gradient problem and its solution during the learning of recurrent neural networks (1998) [30]. Thai-FAQ ChatBots powered by RNN-LSTM (2018) [51]. A knowledge-based ChatBot with effective attention mechanisms based on dual Wasserstein generative adversarial networks (2020) [52].	**Recurrent Neural Network**	R.N.N.s enable data durability [30]. R.N.N.s have been widely used in a range of applications, including language translation, voice recognition modeling, and picture captioning [30]. Experiments demonstrated that the ChatBots equipped with R.N.N. was capable of recognizing and answering	The unmodified R.N.N. is not widely employed due to the vanishing gradient problem [30]. The computation is slow due to its recurring nature [52]. Training R.N.N. models can be challenging [52].	F.A.Q. ChatBots [51] CHARLIE, Proprofs [52].

		queries with a high degree of accuracy [51].		
A search space odyssey using L.S.T.M. (2017) [31]. Thai-FAQ ChatBots powered by RNN-LSTM (2018) [51]. Using a Recurrent Neural Network, Multi-Sensor Guided Hand Gesture Recognition for a Tele-operated Robot (2021) [53].	**Long Short Term Memory (LSTM)**	L.S.T.M. algorithm comes into play, resolving the vanishing gradient problem [51]. Long-term dependency is eliminated using L.S.T.M [51]. Introduce the concept of gates which helps for updating the network's state [31].	All L.S.T.M.s do not all have the same structure [31]. Dropout is much harder to implement in L.S.T.M.s [53]. L.S.T.M.s require more memory to train [53].	Deep probe, Superagent [54].
Neural network-based sequence-to-sequence learning (2020) [33]. A conversational neural model for Chatbots on a smaller scale (2019) [55]. Seq2Seq AI Chatbots with Attention Mechanism (2020) [56]. Seq2seq-model-in-machine-learning (2021) [57].	***Seq2seq***	Two R.N.N.s can be employed as the encoder and decoder, which is the model's most fundamental and original configuration [33]. The Sequence-to-Sequence technique aids in the modeling of lengthier phrases [55]. It is utilized for a wide variety of purposes, including picture captioning, conversational modeling, and text summarization [56].	The most perplexing difficulty is that the model is incapable of handling variable-length sequences [56]. The next one is the size of the vocabulary. If the decoder needs to run through a huge vocabulary, the training process will be slowed [56].	A.L.E.X.A., Google Home [57].
A more detailed attentional neural dialogue model (2016) [34]. Models of ChatBots based on deep learning (2017) [58].	***In-depth Seq2seq***	A model with deeper layers is predicted to perform better [34]. Deep Seq2seq models with several L.S.T.M. layers [58]. Assists in preserving the conversation's long-term context and creates extended outputs [59].	Less interpretable, and hard to debug [60]. Difficult to control: can't specify rules or guidelines [60].	Facebook, Amazon, Slack, Messenger, Zapier, etc. [61].

VII. COMPARATIVE ANALYSIS

In this study, we compare ChatBot techniques under the global metrics which are: usability, performance, speed, technical issues, intelligence, appropriateness of responses, speed of response, repetitiveness, linguistic accuracy, and error management.

A. Classical Approaches Comparison

Classical approaches are also referred to as rule-based techniques because they create replies based on specified rules, which distinguishes them from traditional procedures. Over time, these requirements have gotten more challenging to meet and more complex to understand. For the classical techniques, the parameters we took for comparison are usability, performance, speed, technical issues, and intelligence as can be seen in table III.

Table III: Comparative Analysis of ChatBots' Latest Techniques

Techniques	Usability	Performance	Speed	Technical issues	Intelligence
Pattern Recognition	Yes	Low	Slow	Yes	No
Parsing	Yes	Low	Slow	Yes	No
Markov Chain Models	Yes	Mid	Average	Yes	No
Ontologies	Yes	Mid	Average	Yes	No
AIML	Yes	High	Fast	Yes	Yes
Chat script	No	Mid	Average	Yes	Yes
RDBMS & SQL	Yes	Mid	Average	Yes	No
Language Illusions	Yes	High	Fast	Yes	Yes

Usability in this context means if a technique is still effective. The technique will receive a 'yes or one', if it is still under development and a 'no or 0', otherwise. Performance refers to the approaches' output, which might be low '1', medium '2', or high '3'. Speed is what determines how quickly a procedure produces results; it might be sluggish '1', average '2', or rapid '3'. If there are technical problems with the technique, it will receive a 1; otherwise, it will receive a 0. This indicates whether the method can get better in the future. The ability to learn on

its own is what intelligence means; if a technique is self-learning, it will receive a 1; otherwise, it will receive a 0, see figure VIII.

Based on the parameters and the results against those parameters that were provided following the benefits and drawbacks of the method. We note that, except for chat script, all traditional marketing methods are still effective. AIML and language illusion perform well in terms of performance, with machine learning as the foundation for these methods. Because they handle the data more quickly, both ways have good speed. All techniques have technical problems because they can all be altered to meet ChatBot's requirements. Finally, AIML, chat script, and language illusion all contain intelligence. Overall, figure IX comparison demonstrates that the best traditional approaches are thought to be AIML and language illusion techniques.

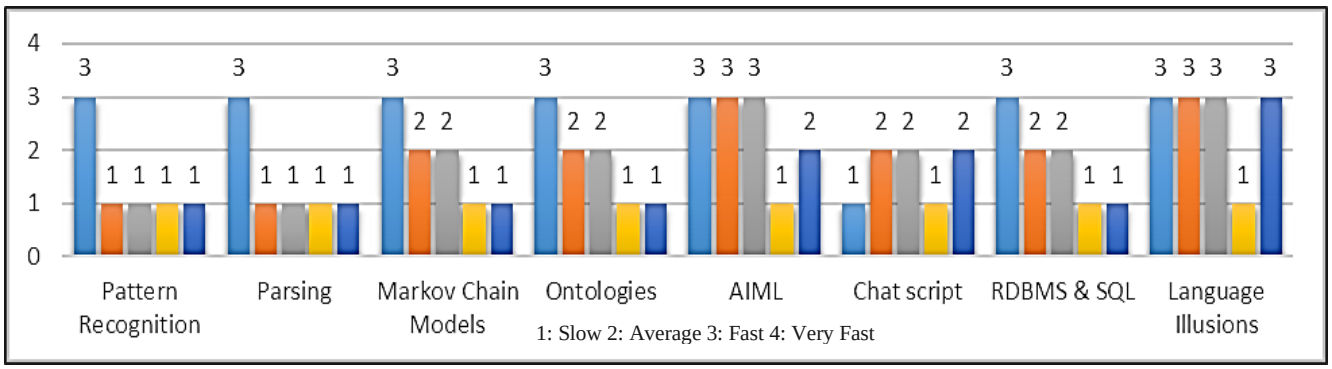

Figure VIII: Comparative Analysis of ChatBots Latest Techniques

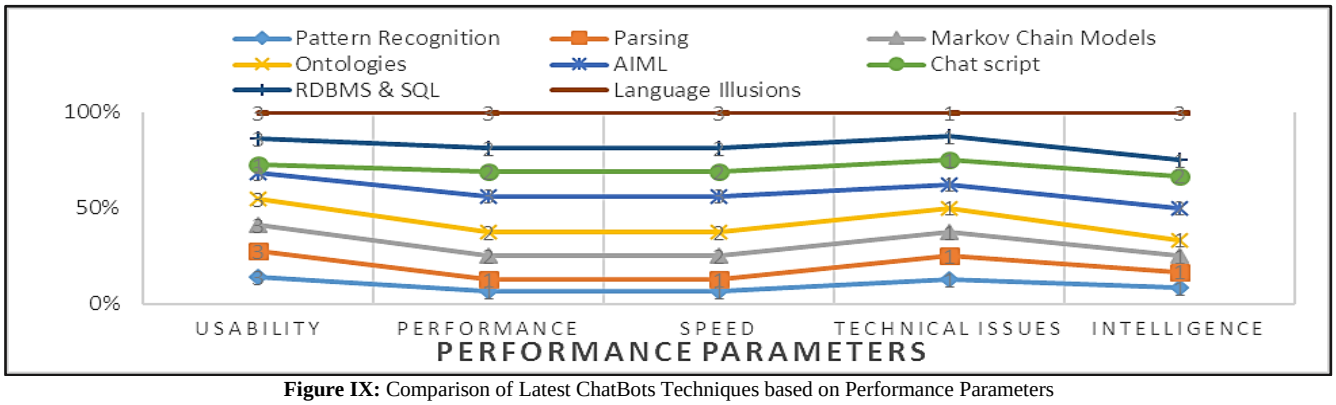

Figure IX: Comparison of Latest ChatBots Techniques based on Performance Parameters

B. Neural-Network-Based Approaches Comparison

The laborious process of formulating rules for every possible utterance-response combination has been replaced with ChatBots powered by neural networks. An artificial neural network can offer a response in one of two ways: either by creating the response from scratch (generative) or by retrieving it from an extensive dataset (retrieval or retrieval-based). The appropriateness of responses, speed of responses, repetitiveness, linguistic accuracy, and error management were used to gauge ChatBot response creation as can be seen in table IV.

Table IV: Comparison of Neural Network-Based ChatBots Approaches

Techniques	Appropriateness of Responses	Speed of Response	Repetitiveness	Linguistic Accuracy	Error Management
Recurrent Neural Network	Low	Low	Yes	Low	No
Long Short Term Memory(LSTM)	Mid	Medium	Yes	Low	No
Seq2seq	High	Fast	Yes	High	Yes
In-depth Seq2seq	Highly	Faster	No	Highly	Yes

Here, "appropriateness of responses" meant that the answer to ChatBot's question was accurate. If the technique yields a common reaction, it will receive a 1; otherwise, a '2', a '3', and a '4' for a very responsive approach. The speed of response refers to how rapidly ChatBot reacts to input. If the speed is slow, it will receive a 1, 2, 3, or 4, depending on how quickly it responds. Repetition determines whether or not the ChatBot repeats its results in response to requests. If the technique is repetitious, it receives a score of 2; otherwise, it gets a score of 1. Linguistic correctness indicates whether a method is producing accurate results or not; it gets a score of 1 for low accuracy, a score of 3 for moderate accuracy, and a score of 4 for highly accurate results. Error management, which comes last, refers to whether a technique has a mechanism to eliminate errors. If a technique can eliminate mistakes, it will receive a score of 1; otherwise, it will receive a score of 3, see figure X.

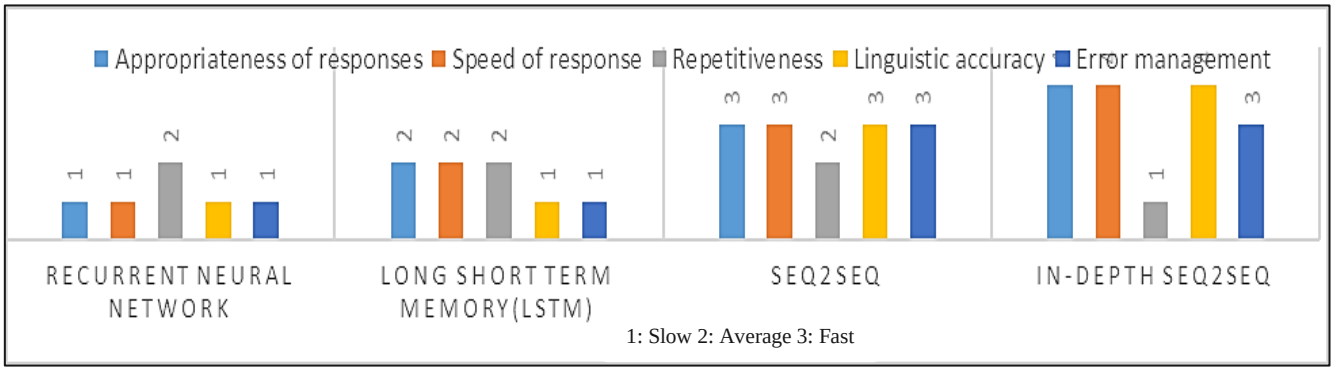

Figure X: Comparison of Neural Network-Based ChatBots Approaches

They are based on the parameters and the results against those parameters that were provided under the benefits and drawbacks of the method. We note that, with the exception of chat script, all traditional marketing methods are still effective. AIML and language illusion perform well in terms of performance, with machine learning as the foundation for these methods. Because they handle the data more quickly, both methods have good speed. All techniques have technical problems because they can all be altered to meet ChatBot's requirements. Finally, AIML, chat script, and language illusion all contain intelligence. Overall, figure XI comparison demonstrates that AIML and language illusion techniques are the best among the traditional approaches.

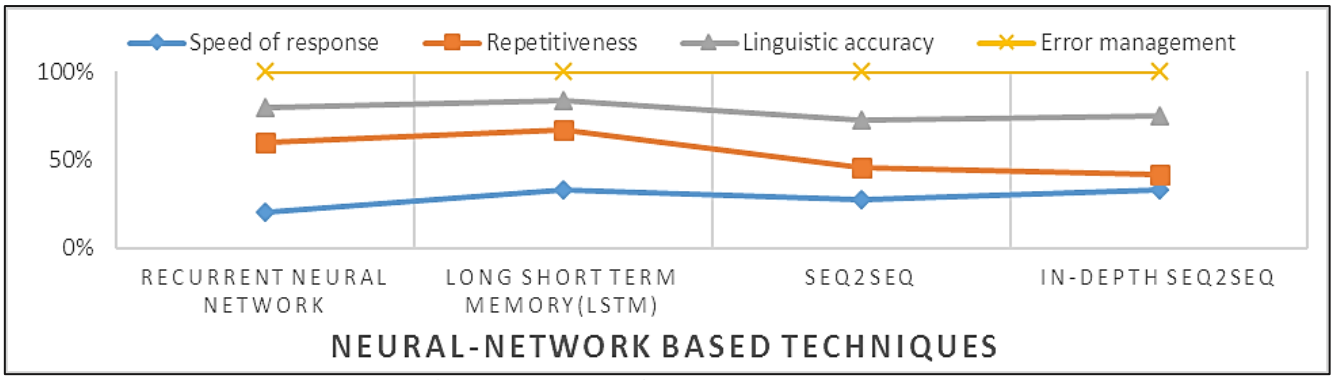

Figure XI: Performance Comparison of Neural Network-Based ChatBots Approaches

VIII. SURVEY ANALYSIS

Beginning with an overview of ChatBots, this article goes on to discuss numerous classic or rule-based strategies as well as neural network-based alternatives in greater depth. The survey article concludes with a table summarizing the most recent findings from field studies. Most recent and important research is looked at, as well as the tools that have been used, areas that could be improved, and the applicability of approaches that have been used. In terms of future development, there are a number of subjects in the field of conversation modeling that have yet to be investigated.

IX. CONCLUSION

Businesses utilize ChatBots to answer frequently asked inquiries and engage customers, according to the company. Human engagement typically occurs when a client is dissatisfied. These ChatBots are being examined to show the different ways to create them. ChatBots are increasingly using neural network-based methods while retaining non-AI benefits. Thus, this review of conversational bot research reflects the approach above. Conversational bots have failed the Turing test in recent years. To evaluate a chatBot's performance, its pros, and cons must be considered. The article concludes that AIML and language illusions are the best rule-based chat techniques worldwide. In neural network-based chatBot development, the in-depth seq2seq approach is optimal. To help academics learn about chatBots, this programme was created.

Acknowledgment

The authors would like to thank almighty Allah first and then the management of, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan, for their support and their assistance throughout this study.

Authors Contributions

The Idea and survey guidance was suggested for this research work by Ahthasham Sajid and Imran Ullah khan along with performance analysis parameters. The Initial draft write-up and English proofreading were done by Junaid Javed. The revision of the draft was improved by Shahnoor and Iqra Tabassum.

Conflict of Interest

The Authors have no conflict of Interest.

Data Availability Statement

The testing data is available in this paper.

Funding

This research received no external funding.

References

[1] Neff, G. (2016). Talking to bots: Symbiotic agency and the case of Tay. *International Journal of Communication, 4915*-31, p. 10.

[2] Güzeldere, G., & Franchi, S. (1995). Dialogues with colorful "personalities" of early AI. *Stanford Humanities Review, 4*(2), 161-169.

[3] Alice Al Foundation. (2023). Ask Alice. Retrieved from: http://www.Alicebot.org/

[4] Turing, A. M. (2009). *Computing machinery and intelligence* (pp. 23-65). Springer Netherlands.

[5] Serban, I. V., Sankar, C., Germain, M., Zhang, S., Lin, Z., Subramanian, S., ... & Bengio, Y. (2017). A deep reinforcement learning chatbot. *arXiv preprint arXiv:1709.02349*, pp. 1-40.

[6] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84-90.

[7] Havrylov, S., & Titov, I. (2017). Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. *Advances in neural information processing systems, 30*.

[8] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

[9] Li, X., Mou, L., Yan, R., & Zhang, M. (2016). Stalematebreaker: A proactive content-introducing approach to automatic human-computer conversation. *arXiv preprint arXiv:1604.04358*.

[10] Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in neural information processing systems, 27*.

[11] Song, Y., Yan, R., Li, X., Zhao, D., & Zhang, M. (2016). Two are better than one: An ensemble of retrieval-and generation-based dialog systems. *arXiv preprint arXiv:1610.07149*.

[12] Clarizia, F., Colace, F., Lombardi, M., Pascale, F., & Santaniello, D. (2018). Chatbot: An education support system for student. In *Cyberspace Safety and Security: 10th International Symposium, CSS 2018, Amalfi, Italy, October 29–31, 2018, Proceedings 10* (pp. 291-302). Springer International Publishing.

[13] Colace, F., De Santo, M., Pascale, F., Lemma, S., & Lombardi, M. (2017). BotWheels: a Petri Net based Chatbot for Recommending Tires. *DATA, 17*, 350-358.

[14] Edwards, B. I., Muniru, I. O., & Cheok, A. D. (2016). Robots to the rescue: a review of studies on differential medical diagnosis employing ontology-based chat bot technology. *Preprints*.

[15] Casillo, M., Clarizia, F., D'Aniello, G., De Santo, M., Lombardi, M., & Santaniello, D. (2020). CHAT-Bot: A cultural heritage aware teller-bot for supporting touristic experiences. *Pattern Recognition Letters, 131*, 234-243.

[16] Bradeško, L., Starc, J., Mladenic, D., Grobelnik, M., & Witbrock, M. (2016, September). Curious cat conversational crowd based and context aware knowledge acquisition chat bot. In *2016 IEEE 8th International Conference on Intelligent Systems (IS)* (pp. 239-252). IEEE.

[17] Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9*(1), 36-45.

[18] Ahmad, S. (2007). Tutorial on natural language processing. *Artificial Intelligence, 810*(161).

[19] Agarwal, R., & Wadhwa, M. (2020). Review of state-of-the-art design techniques for chatbots. *SN Computer Science, 1*(5), 246.

[20] Bradeško, L., & Mladenić, D. (2012, October). A survey of chatbot systems through a loebner prize competition. In *Proceedings of Slovenian language technologies society eighth conference of language technologies* (Vol. 2, pp. 34-37). sn.

[21] Hutchens, J. L., & Alder, M. D. (1998). Introducing megahal. In *New Methods in Language Processing and Computational Natural Language Learning*.

[22] Noy, N. F., & McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology. Stanford: Stanford Knowl Syst Lab, p. 25.

[23] Lenat, D. B. (1995). CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM, 38*(11), 33-38.

[24] Al-Zubaide, H., & Issa, A. A. (2011, November). Ontbot: Ontology based chatbot. In *International Symposium on Innovations in Information and Communications Technology* (pp. 7-12). IEEE.

[25] Marietto, M. D. G. B., de Aguiar, R. V., Barbosa, G. D. O., Botelho, W. T., Pimentel, E., França, R. D. S., & da Silva, V. L. (2013). Artificial intelligence markup language: a brief tutorial. *arXiv preprint arXiv:1307.3091*.

[26] Epstein, R., Roberts, G., & Beber, G. (Eds.). (2009). *Parsing the Turing test* (pp. 978-1). Dordrecht: Springer Netherlands.

[27] Wilcox, B., Wilcox, S., Psych, B. A., & Arts, D. F. (2010). Suzette, the most human computer. *Agent's Processing, Cognition: https://www. chatbots. org/images/uploads/research_papers/9491. p df*.

[28] Lokman, A. S., & Zain, J. M. (2010). Extension and prerequisite: An algorithm to enable relations between responses in chatbot technology. *Journal of Computer Science, 6*(10), 1212.

[29] Bradeško, L., & Mladenić, D. (2012, October). A survey of chatbot systems through a loebner prize competition. In *Proceedings of Slovenian language technologies society eighth conference of language technologies* (Vol. 2, pp. 34-37). sn.

[30] Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 6*(02), 107-116.

[31] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A search space odyssey. *IEEE transactions on neural networks and learning systems, 28*(10), 2222-2232.

[32] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

[33] Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in neural information processing systems, 27*.

[34] Yao, K., Peng, B., Zweig, G., & Wong, K. F. (2016). An attentional neural conversation model with improved specificity. *arXiv preprint arXiv:1606.01292*.

[35] Setiaji, B., & Wibowo, F. W. (2016, January). Chatbot using a knowledge in database: human-to-machine conversation modeling. In *2016 7th international conference on intelligent systems, modelling and simulation (ISMS)* (pp. 72-77). IEEE.

[36] Abdul-Kader, S. A., & Woods, J. C. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications, 6*(7).

[37] Mladenić, D., & Bradeško, L. (2012). A survey of chabot system through a Loebner prize competition. *Proceedings of Slovenian Artificial Intelligence Laboratory, Jozef Stefan Institute in Ljubljana*.

[38] Sharma, S. (2022). What Are The Advantages & Disadvantages of Chatbot in Your Business? Retrieved from: https://allusefulinfo.com/?s=what+are+advantages+disadvantages+of+Chatbots+in+business

[39] Marietto, M. D. G. B., de Aguiar, R. V., Barbosa, G. D. O., Botelho, W. T., Pimentel, E., França, R. D. S., & da Silva, V. L. (2013). Artificial intelligence markup language: a brief tutorial. *arXiv preprint arXiv:1307.3091*.

[40] Arsovski, S., Cheok, A. D., Idris, M., & Raffur, M. R. B. A. (2017, September). Analysis of the chatbot open source languages aiml and chatscript: A review. In *9th DQM International Conference on life cycle engineering and management.*

[41] Wallace, R. S. (2009). *The anatomy of ALICE* (pp. 181-210). Springer Netherlands.

[42] Yan, R., Song, Y., & Wu, H. (2016, July). Learning to respond with deep neural networks for retrieval-based human-computer

conversation system. *In Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 55-64).

[43] Caldarini, G., Jaf, S., & McGarry, K. (2022). A literature survey of recent advances in chatbots. *Information*, *13*(1), 41.

[44] Satu, M. S., & Parvez, M. H. (2015, November). Review of integrated applications with aiml based chatbot. In *2015 International Conference on Computer and Information Engineering (ICCIE)* (pp. 87-90). IEEE.

[45] Satu, M. S., Parvez, M. H., & Mamun, S. A. (2015).Review_of_integrated_applications_with_AIML_based_Chatbots. Retrieved from: https://www.researchgate.net/profile/Md-Satu/publication/304063740_Review_of_integrated_applications_with_AIML_based_chatbot/links/5765418a08aedbc345f37e4e/Review-of-integrated-applications-with-AIML-based-chatbot.pdf

[46] Wilcox, B., & Wilcox, S. (2014). Making it real: Loebner-winning chatbot design. *ARBOR Ciencia, Pensamiento y Cultura*, *189*(764), 1-13.

[47] Ramesh, K., Ravishankaran, S., Joshi, A., & Chandrasekaran, K. (2017, October). A survey of design techniques for conversational agents. In *Information, Communication and Computing Technology: Second International Conference, ICICCT 2017, New Delhi, India, May 13, 2017, Revised Selected Papers* (pp. 336-350). Singapore: Springer Singapore.

[48] Bradeško, L., & Mladenić, D. (2012, October). A survey of chatbot systems through a loebner prize competition. In *Proceedings of Slovenian language technologies society eighth conference of language technologies* (Vol. 2, pp. 34-37). sn.

[49] Ahmad, N. A., Che, M. H., Zainal, A., Abd Rauf, M. F., & Adnan, Z. (2018). Review of chatbots design techniques. *International Journal of Computer Applications*, *181*(8), 7-10.

[50] Ehrenbrink, P., Osman, S., & Möller, S. (2017, November). Google now is for the extraverted, Cortana for the introverted: investigating the influence of personality on IPA preference. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction* (pp. 257-265).

[51] Muangkammuen, P., Intiruk, N., & Saikaew, K. R. (2018, November). Automated thai-faq chatbot using rnn-lstm. In *2018 22nd International Computer Science and Engineering Conference (ICSEC)* (pp. 1-4). IEEE.

[52] Kim, S., Kwon, O. W., & Kim, H. (2020). Knowledge-grounded chatbot based on dual wasserstein generative adversarial networks with effective attention mechanisms. *Applied Sciences*, *10*(9), 3335.

[53] Qi, W., Ovur, S. E., Li, Z., Marzullo, A., & Song, R. (2021). Multi-sensor guided hand gesture recognition for a teleoperated robot using a recurrent neural network. *IEEE Robotics and Automation Letters*, *6*(3), 6039-6045.

[54] Patil, S., Mudaliar, V. M., Kamat, P., & Gite, S. (2020). LSTM based Ensemble Network to enhance the learning of long-term dependencies in chatbot. *International Journal for Simulation and Multidisciplinary Design Optimization*, *11*, 25.

[55] Mathur, S., & Lopez, D. (2019). A scaled-down neural conversational model for chatbots. *Concurrency and Computation: Practice and Experience*, *31*(10), e4761.

[56] Sojasingarayar, A. (2020). Seq2seq ai chatbot with attention mechanism. *arXiv preprint arXiv:2006.02767*.

[57] geeksforgeeks . (2022). seq2seq-model-in-machine-learning. Retrieved from: https://www.geeksforgeeks.org/seq2seq-model-in-machine-learning/

[58] Csaky, R. (2019). Deep learning based chatbot models. *arXiv preprint arXiv:1908.08835*.

[59] Serban, I., Sordoni, A., Lowe, R., Charlin, L., Pineau, J., Courville, A., & Bengio, Y. (2017, February). A hierarchical latent variable encoder-decoder model for generating dialogues. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 31, No. 1).

[60] Github. (2022). Lecture 8 – Translation, Seq2Seq, Attention. Retrieved from: https://x-wei.github.io/xcs224n-lecture8.html#nmt-advantages-disadvantages_1.

[61] softwaretestinghelp . (2023). Top 12 Best AI Chatbots For 2023 [Review & Comparison]. Retrieved from: https://www.softwaretestinghelp.com/best-ai-Chatbots/

[62] Madhumitha, S., Keerthana, B., & Hemalatha, B. (2020). Interactive Chatbots Using AIML. Retrieved from: https://www.ijana.in/papers/62.pdf

[63] Colace, F., De Santo, M., Lombardi, M., Pascale, F., Pietrosanto, A., & Lemma, S. (2018). Chatbot for e-learning: A case of study. *International Journal of Mechanical Engineering and Robotics Research*, *7*(5), 528-533.

[64] The Medical Futurist. (2021). The Top 12 Health Chatbots. The Medical Futurist , p. 106–111. Retrieved from: https://medicalfuturist.com/top-12-health-chatbots/

[65] Venkat, k. (2020). Examples and Best use cases of chatbots in Banking Industry. Retrieved from: https://medium.com/voice-tech-podcast/examples-and-best-use-cases-of-chatbots-in-banking-industry-1a46cebdb0cd