Optimizing Customer Interactions: A BERT and Reinforcement Learning Hybrid Approach to Chatbot Development

Dr. K.R.Praneeth[1], Dr. Taranpreet Singh Ruprah[2], Dr. J Naga Madhuri[3], Dr. A L Sreenivasulu[4],
Syed Shareefunnisa[5], Dr. Vuda Sreenivasa Rao[6]

Assistant Professor, School of Management, The Apollo University, The Apollo Knowledge City Campus,
Chittoor, Andhra Pradesh, India[1]
Associate Professor, Amity University, Jaipur, Rajasthan, India[2]
Assistant Professor, Department of English, Velagapudi Rama Krishna Siddhartha Engineering College
(Deemed to be University), Kanuru, Vijayawada, India[3]
Professor of CSE, Vignana Bharathi Institute of Technology, Telangana, Aushapur(V), Ghatkesar(M) Medchal (D),
Hyderabad, Telangana, India[4]
Assistant Professor, VFSTR Deemed to be University, Guntur, Andhra Pradesh, India[5]
Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation,
Vaddeswaram, AP, India[6]

Abstract—**In the case of chatbots massive progress has been made, but problems remain in handling the complexity of the sentence and in context relevance. Traditional models can be rather insufficient when it comes to providing various levels of detail in the responses to the end-users' questions, particularly when referring to customer support scenarios. To overcome these limitations, this research comes up with a new model which combines the BERT model with DRL. Through DRL, BERT pre-training is adding flexibility and correspondence to correctly perceive contextual delicate matters in the response. The proposed method includes the following pipeline where in; data tokenization, conversion to lowercase characters, lemmatization and then passes through the BERT fine-tuned model. DRL is utilized to optimize the chatbot's response in the light of long term rewards and the conversational history, the interactions are formulated as a Markov Decision Process with the reward functions based on cosine similarity of the consecutive responses. This makes it feasible for the chatbot to provide context based replies in addition to the option of constant learning for enhanced performance. It also proved that the accuracy and relevance of the BERT-DRL hybrid system were higher than traditional models according to the BLEU and ROUGE scores. The performance of the chatbot also increases with the length of the conversation and the transitions from one response to the other are coherent. This research contributes to the field through the integration of BERT in understanding language and DRL in the iterative learning process in the innovation within the flaws of chatbot technologies and establishing a new benchmark for conversational AI in customer service settings.**

Keywords—Chatbots; BERT (Bidirectional Encoder Representations from Transformers); RL (Reinforcement Learning); customer service; responsiveness

I. INTRODUCTION

A chatbot is a computer program that talks to people like a human. It can use text or speech to have conversations with users. In customer service, chatbots are really important in making customer support better and more efficient. These smart systems are used on websites, mobile apps, and messaging apps to help customers quickly with their questions. Chatbots use special tools to understand what users are saying, figure out what they want, and give them the right answers. They can do many different jobs, like answering questions, giving product information, helping with problems, and doing transactions [1]. Chatbots help make customers happy by answering quickly and being available all the time. This also helps businesses work better and faster. But, chatbots need to understand what users are saying, have good conversations, and adjust to different ways people talk and different situations in order to work well. Ongoing research and improvements in NLP, machine learning, and dialogue management are very important. They help make customer service chatbots better and make sure they can talk to users without any problems [2].

Chatbots are helpful in customer service because they can quickly and accurately respond to many different questions. By using smart computer programs, chatbots can understand what users want, get important information, and quickly give the right answers [3]. This helps businesses grow their customer support without sacrificing quality or speed, so they can use their resources better and work more efficiently. Additionally, chatbots help make customers feel more connected and loyal by giving them personalized experiences based on what they like and how they act. By using data analysis and machine learning, chatbots can use past conversations and customer information to predict what customers might need, suggest helpful products or services, and solve problems before they get worse [4]. Chatbots are important for businesses because they help to keep customers happy and coming back [5]. They do this by building good relationships and having positive conversations with people. Yet, to make chatbots work well in customer service, it's important to plan carefully, keep improving them, and

follow the best ways of doing things. Businesses need to spend money on the right technology, talented people, and good ways of working to make their chatbots work as well as possible. This includes making them easy for people to use, creating good conversations for them, teaching them how to learn, and keeping an eye on how well they are doing. Moreover, as customers' needs change, chatbots need to change too. They need to keep up with the latest trends, preferences, and technologies to stay useful in the changing world of customer service [6].

Being quick and accurate when talking to a chatbot is really important. It helps make sure users have a good experience and that businesses and customers get what they want. Responsiveness means the chatbot can quickly answer the questions and requests from users. In the fast-paced digital world of today, people anticipate prompt assistance and solutions for their problems. A chatbot that responds quickly not only meets these expectations but also helps to make users less frustrated, keeps them interested, and makes them feel like they can trust it [7]. In the same way, it's really important for the chatbot to give accurate and helpful responses that match what the user needs. A good chatbot understands what users mean, understands their questions in the right way, and gives the right information or takes the right actions. Giving wrong answers can make users feel confused, misinformed, and unhappy. This can make the chatbot seem less trustworthy and helpful. Additionally, giving wrong answers can make users feel frustrated and uninterested, which can make them think negatively about the brand or service [8].

Research wants to use a combination of BERT and reinforcement learning to make chatbots better. Traditional chatbots have limitations, and research want to improve how quick and accurate they are. BERT is really good at understanding language and making relevant responses. By using BERT's smart understanding of words and meanings, chatbots can be better at understanding what users want and giving them the right answers. Alternatively, reinforcement learning is a useful way to improve how chatbots manage conversations and make decisions [9]. By teaching the chatbot by trying different things and learning from user feedback, reinforcement learning helps the system improve its responses over time. This smart way of learning helps the chatbot get better and better at talking to people over time. It gets faster and more effective at having conversations in real life. By using a mix of BERT and reinforcement learning, chatbots can better understand and respond using words and also improve how they manage conversations. This mixed method makes the chatbot better at understanding and answering questions from users. It also helps the chatbot to learn and change according to what users want and how they talk. Overall, the hybrid system offers a positive solution for developing smarter and more effective chatbots that can improve user experiences in various contexts, such as customer service.

The primary contribution of the work is as follows:

- Making of a new framework of the chatbot that integrates the ability of BERT to consider a context along with the reinforcement learning to improve the quality of the chatbot's answers with time based on the feedback from the user.

- Substantial increase in the quality and pertinence of chatbot answers, as it was proven by higher BLEU and ROUGE scores in comparison to a basic chatbot model.

- Use of the progressive learning model in which the chatbot adjusts its response to the users depending on the results obtained in the process.

- Enhanced customer satisfaction and interactivity as concomitant with presenting actual and highly and closely related responses to specific queries, thereby making customer services more fit into the natural context.

- Establishment of a general framework of a chatbot that can be easily fitted into other domains, thus, can be applicable to a wide range of client services.

The order of the remaining sections is as follows. An introduction is given in Section I. The literary sections are shown in Section II. This is the issue within the conventional methods provided in Section III. The suggested methodology for the study's research design, data gathering strategies, and analytic approaches are covered in Section IV. The efficacy measures are presented, and the results are compiled in Section V. Section VI delivers more studies and a conclusion.

II. RELATED WORK

Dhyani and Kumar [9] explains how to use deep learning to make a better Chatbot. Implementing the Neural Machine Translation (NMT) model using the TensorFlow software library. Learning and gathering information for creating a model is a very important but challenging task. Bidirectional Recurrent Neural networks with attention layers are employed to improve responses for lengthy or wordy sentences. The information used to teach the computer model in the paper comes from Reddit. The model was made to translate English to English. This work aims to make the model more confused and learn more quickly and to measure the translation quality using Bleu Score for the same language. Research did experiments with TensorFlow using python 3. 6 The confusion, learning speed, language evaluation score and average time for every 1000 steps are 56. 10, 00001, 3016 and 45 One period is finished after taking 23,000 steps. The paper also looks at how MacBook Air can be used for neural network and deep learning. In the future, research will also make a healthcare Chatbot to help patients with diseases like COVID-19, diabetes, high blood pressure, and heart problems. by giving details about the disease, suggesting what foods to eat, and explaining how to handle emergencies.

A chatbot is a computer program that talks like a person and can have a conversation with a human. One important job in artificial intelligence and understanding human language is to study how people talk to each other. Since artificial intelligence started, it has been really difficult to make a good chatbot. Even though chatbots can do many different things, their main job is to understand what people say and reply in a helpful way. In the past, chatbot designs were made using basic statistics or written rules and templates. Around 2015, end-to-end neural networks

replaced other models because they can learn better. Right now, the encoder-decoder recurrent model is very important for modeling conversations. This design comes from the field of machine translation using neural networks and it worked very effectively in that area. So far, many new features and changes have made chatbots better at talking to people. In this paper, research studied a lot of new books and articles. Research looked at a lot of articles from the past five years about chatbots. Next, research talked about other research on the topic and the AI ideas needed to make a smart chatbot using deep learning. Then, research showed a plan for creating a helpful chatbot for healthcare. In the future, research will be better at spotting and diagnosing bots, as research will have more advanced tools to help us understand their symptoms, like how intense they are, how long they last, where they are happening, and a more detailed description of what is going on [10].

AI, ML, and NLP are changing the way organizations do things. As more data comes in and AI systems get better at using it to help businesses, people are getting more excited about AI. Large amounts of data, computer power, better ways to solve problems, easy-to-use tools, and systems have made companies use AI to make their business better and make more money. These technologies help all kinds of businesses, from farming to finance. AI, ML, and NLP are helping companies with things like customer service, predicting what might happen, making things more personal for customers, recognizing pictures, understanding emotions, and working with documents online and offline. This study had two main goals. First, research look at how AI is used in business. Then research check if these uses make customers more loyal by looking at data from 910 different companies. The data includes scores for four different AI features: AI customer service, predictive modeling, personalized machine learning, and natural language processing integration. The goal is to measure how loyal customers are using a simple yes or no answer. All the qualities are rated on a scale of 1 to 5. Research used six different types of computer programs to help us learn from and make predictions about data. These were Logistic regression, KNN, SVM, Decision Tree, Random Forest, and Ada boost Classifiers. The abilities of the different algorithms were tested using confusion matrices and ROC curves. The decision tree had an accuracy of 0. 532 and KNN had an accuracy of 0. 570This study shows that businesses can use AI, ML, and NLP to look at data and find important information. This can help them automate tasks and plan business strategies. In order to stay competitive and retain customer loyalty, companies should begin incorporating them into their business strategies [11].

It is important to understand how people feel when they interact with robots as customers. This can be seen from the reviews they share online. Knowing this is important for predicting whether people will want to use service robots in the future. Qualitative analysis gives us lots of helpful information from data, but it takes lots of time and effort to do. Experts have talked about how helpful it is to use algorithms to tell the difference between different emotions. This study looks at the good and bad sides of using qualitative analysis and machine learning methods together, using both human and machine intelligence. Research took 9,707 customer reviews from two big social media sites (Ctrip and TripAdvisor). The reviews

were about 412 hotels in 8 different countries. The study found that customers really like service robots, and they feel happy, amazed, and excited when they interact with them. Customers are not happy when service robots don't work and they can't use them. Robots that help people can make them feel more emotions when they move. The results also show how different cultures can affect how customers feel about service robots. The research shows that using a mix of different methods can make machine learning faster and more efficient. This aids in addressing certain issues with machine learning, such as the difficulty in comprehending concepts and the limited range of emotions available [12].

Analyzing customer feelings is very important for understanding how customers feel about products, especially in online stores where there are a lot of customer reviews. The changes in e-commerce reviews, like adding pictures, videos, and emojis, makes it more complex to analyze how people feel about products. Old-fashioned text models might have a hard time understanding feelings shown in things that are not written. This paper suggests a better way to understand how people feel about products on online stores, to make it better for shoppers. The plan involves using Fejer Kernel filtering to estimate data points in the E-commerce dataset. Research use a fuzzy dictionary to find important words in the E-Commerce dataset. The information research used for the study was collected using a method called Optimized Stimulated Annealing to pick out the most important features. Customer opinions are sorted using the BERT deep learning model. The model gives us the opinions of consumers in the E-Commerce dataset. The way customers feel about products in the E-commerce data decides how they are grouped. The test showed that the new model is better at correctly categorizing things on the online shopping site. This study helps to improve how research understand and use sentiment analysis for online shopping. It is a milestone in the development of more perceptive systems that can understand and respond to customer input [13].

III. PROBLEM STATEMENT

Existing customer service chatbots regularly face several limitations that the proposed development of a Responsive Customer Service Chatbot using a BERT and RL hybrid system objectives to address. Traditional chatbots commonly depend on rule-based totally or statistical fashions that may conflict with information context, dealing with complicated queries, and generating coherent responses [12]. They can also fail to evolve to evolving consumer options and remarks, ensuing in a static person enjoy [11]. Also, existing models often lack customized interplay abilities and can exhibit limited mastering from user interactions because of insufficient training data and simplistic algorithms. These chatbots normally do now not leverage superior language expertise techniques, main to suboptimal handling of nuanced language and conversational context. The proposed hybrid system, combining BERT's deep contextual embeddings with RL's adaptive learning mechanisms, goals to overcome these obstacles through improving the chatbot's capability to apprehend and generate contextually relevant responses.

IV. PROPOSED RESPONSIVE CUSTOMER SERVICE CHATBOT USING A BERT AND REINFORCEMENT LEARNING HYBRID SYSTEM

The proposed method for enhancing the performance of chatbots and virtual assistants in comprehending and responding to human language involves a systematic pipeline. Firstly, data is collected, followed by pre-processing steps like tokenization, lowercasing, removing stop words, and lemmatization to clean and prepare the data. Subsequently, the processed data is fed into a BERT model, renowned for its ability to understand contextual nuances in conversations, thus providing dialogue state representation. Finally, Deep Reinforcement Learning (DRL) is employed to optimize the relevance of responses generated by the BERT model, ensuring that the chatbot or virtual assistant delivers responses that are contextually appropriate and accurate. This comprehensive approach not only improves the understanding of conversation context but also ensures the generation of responses that are highly relevant to user queries or statements, thereby enhancing the overall performance of the conversational system. This is visually presented in Fig. 1.
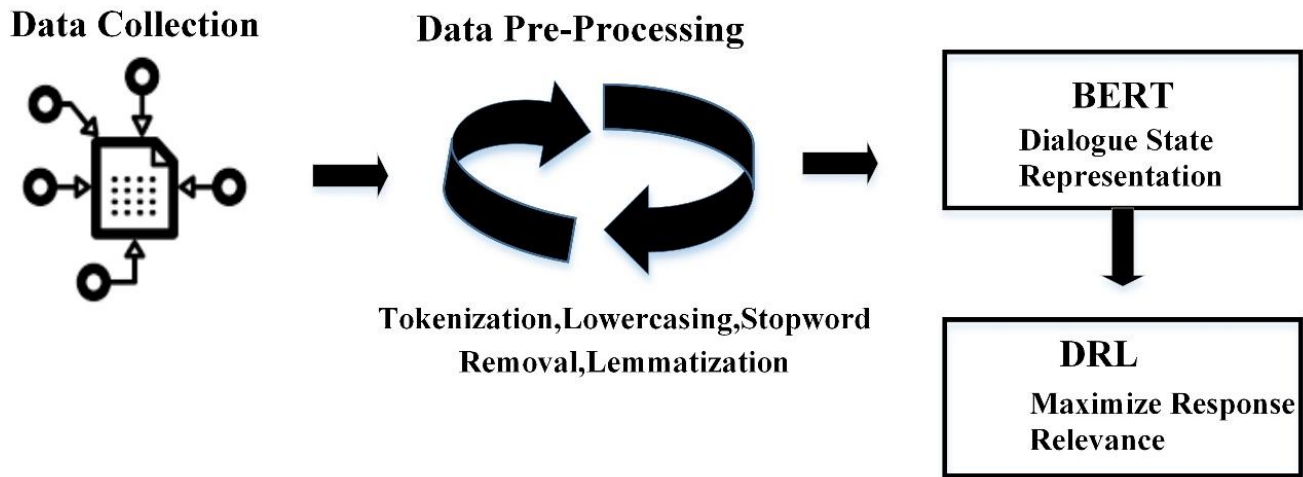

Fig. 1. Proposed method.

A. Data Collection

Customer service chatbot data collected from Kaggle refers to a dataset sourced from the Kaggle platform, which contains conversational data typically used to train and evaluate customer service chatbots. This dataset usually includes text exchanges between customers and customer service representatives across various domains, such as retail, telecommunications, or technology, and serves as valuable input for developing and fine-tuning natural language processing models aimed at automating customer support interactions[10]. The dataset incorporates a cause tag for "good-bye," which encompasses styles like "Bye," "See you later," and "Goodbye," and is associated with responses which includes "See you later, thanks for travelling!" and "Have a nice day!" This established layout facilitates in training the chatbot to recognize and reply appropriately to one of a kind person interaction, ensuring that it could deal with common conversational exchanges effectively.

B. Pre-processing

1) Tokenization: Breaking down sentences into individual words or tokens allows for better analysis of the text. It helps in understanding the structure of the sentences and facilitates further processing steps.

2) Lowercasing: Converting all text to lowercase ensures consistency in the dataset, preventing the model from treating words with different cases as different entities. This step also reduces the vocabulary size and improves generalization.

3) Stopword Removal: Eliminating common stopwords such as "and", "the", "is", etc., helps in reducing noise in the dataset. Since these words occur frequently but often carry little semantic meaning, removing them can improve the quality of the text data.

4) Lemmatization: Reducing words to their base or dictionary form through lemmatization helps in normalizing the text and reducing dimensionality. It ensures that different forms of the same word are treated as one, thus improving the efficiency of downstream tasks like sentiment analysis or topic modelling [11].

C. BERT

In customer support chatbots, BERT's role is essential in understanding user queries and generating relevant responses. Through means of users interact with the chatbot, their queries undergo tokenization and are fed into the BERT model for natural language understanding. Bidirectional nature captures complicated word relationships, decoding the nuanced context within the user's message. This is essential user's need, whether it involves in search of information, reporting an issue, or providing feedback. Once the user's intent is classified, the chatbot employs task-specific output layers to customize its responses. For instance, if a user seeks assistance with a

technical issue, the chatbot may direct the query to a troubleshooting module trained specifically for such concerns. Similarly, if the user expresses dissatisfaction or provides feedback, the chatbot may route the query to a sentiment analysis module to assess customer sentiment and respond accordingly. Fig. 2 illustrates the framework of a pre-trained BERT model, highlighting its architecture and key components for natural language processing tasks.

Prior to implementation, the BERT model is a fine tune to domain-specific data on customer support services. This design fine-tuning ensures that the model adapts to complex customer enquiries and support interactions, improves the ability to provide accurate and helpful information and fine-tunes the model templates for tasks such as issue resolution, queries and sentiment analysis, thereby helping users maximize its effectiveness. The cross-entropy loss used during fine-tuning is represented in Eq. (1):

$$Loss: \sum Loss = \sum_i Y_{i\log(\widehat{Y_I})} \qquad (1)$$

Using thorough express ratings or implicit indicators that indicate user involvement, the chatbot uses feedback to alter parameters and increase its ability to serve customers effectively. In simple terms, BERT serves as the foundation for customer service chatbots, allowing them to recognize user requests, classify intentions, generate contextually appropriate answers, and continuously enhance via comments. Utilizing BERT's simultaneous understanding and first-rate tuning capabilities, interaction with chatbots may provide personalized, effective support, enhancing the user experiences and increasing client pride is given in Eq. (2)

Feedback Loop:

$$Update: \theta_{BERT=}\theta_{BERT-\alpha \ \nabla\theta_{BERT}Loss} \qquad (2)$$

The following formula depicts each parameter updating step in the continuous learning process, where α is the development rate and ∇θ_BERT signifies the change in gradient in relation to the BERT model parameters. Words in sentences are masked in BERT during pre-training; by predicting them, BERT learns bidirectional representations rather than reading text in a one-pass manner as most models do. Through having an interchange between the two directions, BERT is able to seize the contextual meaning of words making it very efficient in the interpretation of queries made in customer service chat bots. In the second step known as the 'fine-tuning' the BERT is trained for certain tasks such as for customer support, by feeding it with relevant data, which helps in providing appropriate responses. This, coupled with the task-specific fine-tuning, makes BERT ideal for this scenario, especially given its bidirectional context, out-competing models that do not have that or are trained in an entirely different model for an entirely different task.
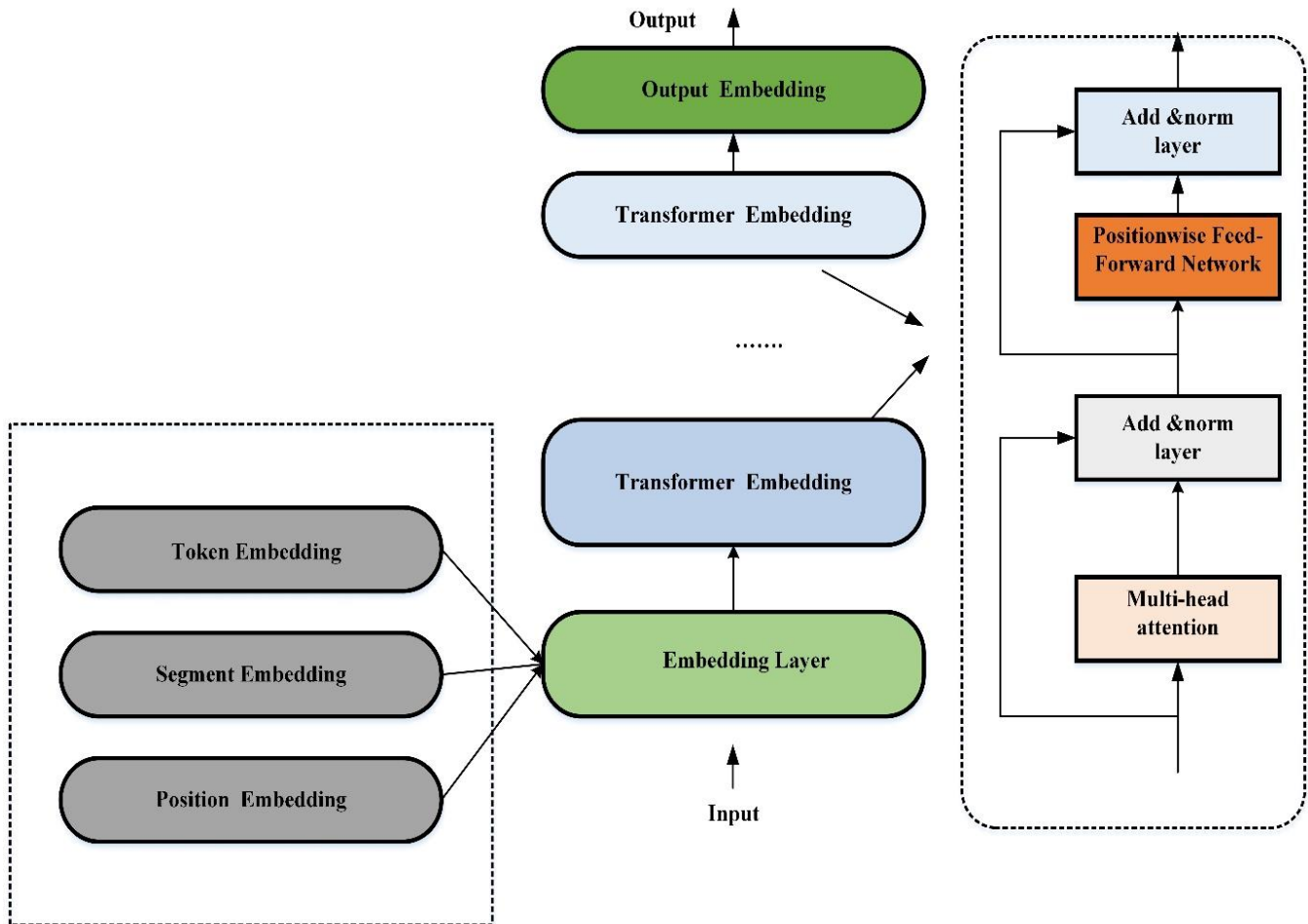

Fig. 2.   Framework of a Pre-trained BERT model.

D. Deep Reinforcement Learning to Maximize Response Relevance

The development of a bidirectional contextual chatbot involves utilizing deep reinforcement learning (RL) to maximize response relevance and effectiveness. The chatbot operates within a Markov Decision Process (MDP), modeling the conversation as a reinforcement learning problem. It consists of states, actions, transition functions, and reward functions. The goal is to learn a policy that maximizes long-term rewards, ensuring responses are not only relevant to the current turn but also consider future discourse implications.

1) Markov Decision Process (MDP): A Markov Decision Process is used for modeling the discussion. MDP consists of states S, actions A, a transition function P, and a reward function R. Provided an MDP (S, A, P, and R), the algorithm is taught to identify an approach that resolves this issue. From an algorithm perspective, policy is a conditional likelihood distributed on the set of activities A. During the encounter, the agent takes action in accordance with the policy. The surrounding area changes state in response to the agent's behavior. Furthermore, the agent communicates with the surroundings at every discontinuous time step (t = 0, 1, 2, ...) is represented in Eq. (3).

$$p(s_{t+1}, r_{t+1} | s_t, a_t) \tag{3}$$

2) Reward Definition: The study suggests two rewards for eliciting targeted replies and achieving specified goals in conversations. It solves the problem of encoder-decoder-based models producing incoherent or insignificant outcomes. Prospective function are defined by sequential rotations and interactions among acts and prior declarations. The reward for every action is represented by r1, and the cosine similarity between actions determines the initial reward at the present state. Let h_t and h_{t+1} the graphical representations obtained from the agents for successive rounds. The cosine similarity among h_t and h_{t+1} yields a preliminary reward at present stage s_t is represented in Eq. (4).

$$r_1 = \cos(h_t, h_{t+1}) = \cos\left(\frac{h_t, h_{t+1}}{||h_t \, h_{t+1}||}\right) \tag{4}$$

3) Conversation Simulation: Although the pre-trained encoder-decoder enables the algorithm to produce coherent responses based on the discussion the past, applying RL improves the model's capacity to produce responses that are optimized for long-term objectives. To replicate the chat as follows. In the initial stage, researchers extract an input sentence via training dataset that includes conversational history as background data and give it to the initial agent. The initial agent encrypts the inputs as a vector c_L and analyzes it to provide an outcome for the following round. The second agent

modifies the simulation's condition by integrating the dialogue record and output. It instantly encodes the altered state into its visual form and interprets it to generate a new response, which is then passed again to the original agent and repeating. At the completion of the simulation, the right-context cR is a sequence of k successive utterances { $s_{t+1}, s_{t+2}, \ldots \ldots s_{t+k},$ } to the correct portion of the produced response s_t. The change in distribution of probabilities, which represents the policy, has been initialized with a pre-trained BERT-based models. Candidate replies have been produced using the above distribution. This is represented in Eq. (5).

$$\pi = \rho_{bert2bert(a'_t)}[[s'_t, c'_l]] \tag{5}$$

The agent's goal is to maximize the expected accumulated reward through a series of actions (responses) during the entire discussion. To accomplish this, the parameters of the model are modified utilizing the policy gradients method. An approach for learning is employed, modeling debate for k turns and using policy gradient approaches to generate parameters that maximize the expected future reward. The purpose is to maximize the predicted cumulative value from the scheduled sequence of actions. The loss function is computed based on an anticipation for tasks and rewards in history. The upward slope of the loss function is generated using the chain rule, and this allows for adjustments to parameters to improve the model's effectiveness [12]. The MDP in the context of a Customer Service chatbot, one can refer to the following example of an interaction. Let a user ask about availability of a certain product or for some recommendations. Based on the state (the user query), the chatbot (agent) has the choice of the next action (response). For example, instead of 'Yes, that is correct,' it might say 'Actually, let me check that for you.' This response leads the conversation flow from one state to another – the user can then ask particular questions and elicit further from the chatbot. Their function is to achieve the maximum total reward that is to offer most helpful and pertinent answers. Cosine similarity is particularly important here: reward functions lie in this context. With reference to the ideal/expected response, cosine similarity makes ensures that the flow of the conversation is based on the similarity index between the chatbot and the user's input. High cosine similarity reveals that the chatbot's response is on the right path, the low one means that the chatbot should improve its response. While the reinforcement learning continues, the conversation patterns of the chatbot get closer to the needs of the users. For instance, at the start it may just generate general responses but it ought to adopt correct contextual replies such as: "The product is in our store at San Jose, shall I set it aside for you?" This enhancement comes from learning through feedback loops on the actions that the bot undertakes. Fig. 3 depicts the response generator utilizing deep reinforcement learning, showcasing its architecture and interaction mechanism for generating conversational responses.
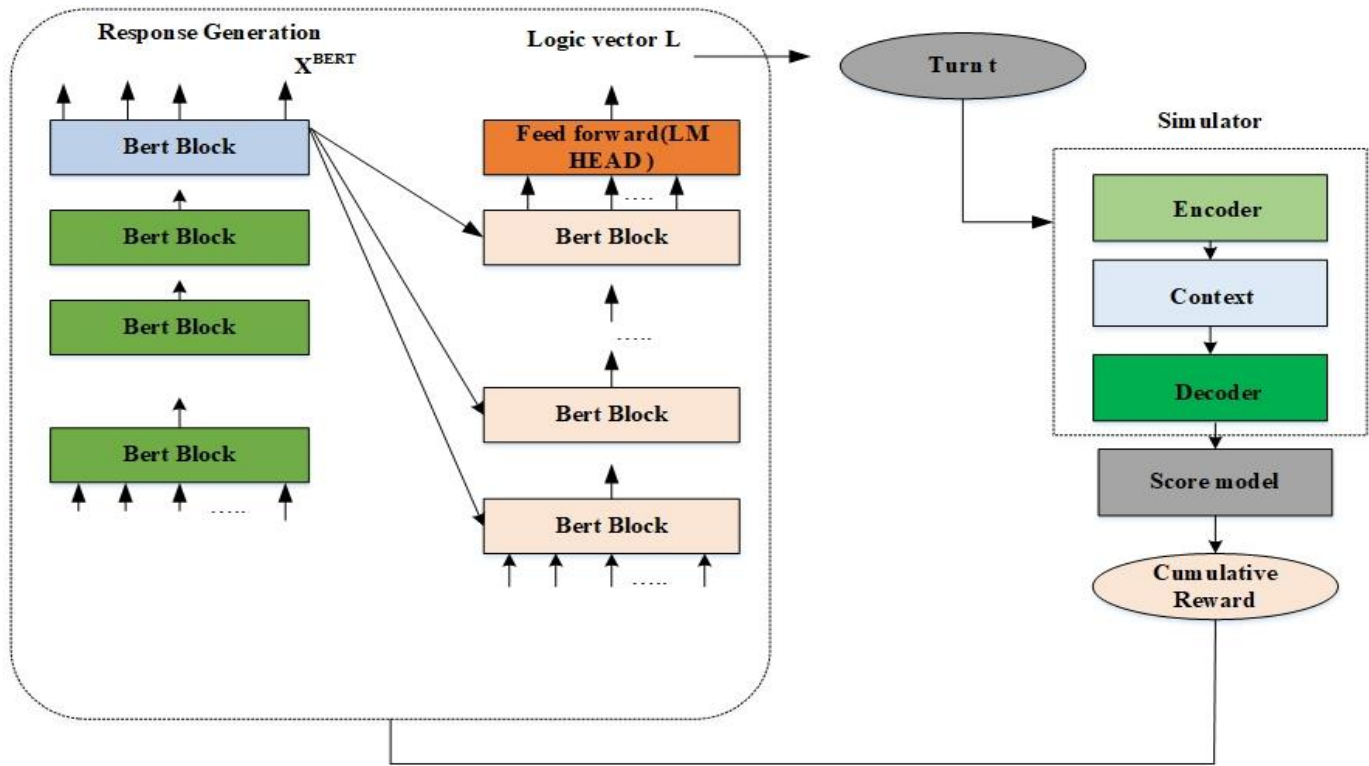
Fig. 3.   Response generator using deep reinforcement learning.

V.   RESULT AND DISCUSSION

The proposed hybrid BERT and Reinforcement Learning chatbot established huge improvements in customer service interactions. The BERT version successfully generated contextually relevant responses, whilst reinforcement getting to know optimized these responses primarily based on actual-time remarks. Evaluation metrics, which include BLEU and ROUGE ratings, showed a marked improvement in response accuracy and relevance compared to conventional fashions. User satisfaction and engagement multiplied, because the chatbot correctly addressed patron queries and tailored over time. The machine's iterative mastering technique resulted in an extra intuitive and responsive chatbot, improving ordinary consumer revel in and proving the efficacy of combining BERT with reinforcement getting to know for dynamic service interactions.

TABLE I.   COMPARSION OF DIFFERENT DEEP LEARNING METHODS

	BLEU 1	BLEU 2	BLEU 3	ROUGE Precision	ROUGE Recall	ROUGE F1 score
CNN	0.433	0.340	0.200	0.318	0.278	0.214
LSTM	0.420	0.355	0.204	0.323	0.311	0.290
GRU	0.477	0.322	0.258	0.389	0.318	0.310
BERT	0.480	0.350	0.246	0.355	0.330	0.370
BERT-RL	0.499	0.399	0.255	0.359	0.340	0.390

Table I   compare the performance of different deep learning methods, including CNN, LSTM, GRU, BERT, and BERT-RL,

in the context of a customer service chatbot. The metrics reported include BLEU scores (BLEU1, BLEU2, BLEU3) for measuring response similarity to human references, and ROUGE scores (ROUGEPrecision, ROUGERecall, ROUGEF1 score) for evaluating response overlap [13] [14]. Generally, BERT-based methods, especially when combined with reinforcement learning (BERT-RL), demonstrate higher BLEU and ROUGE scores compared to traditional RNN-based architectures like LSTM and GRU, as well as CNN-based models, indicating their superior performance in generating contextually relevant and accurate responses in customer service interactions. This is visually represented in Fig. 4.
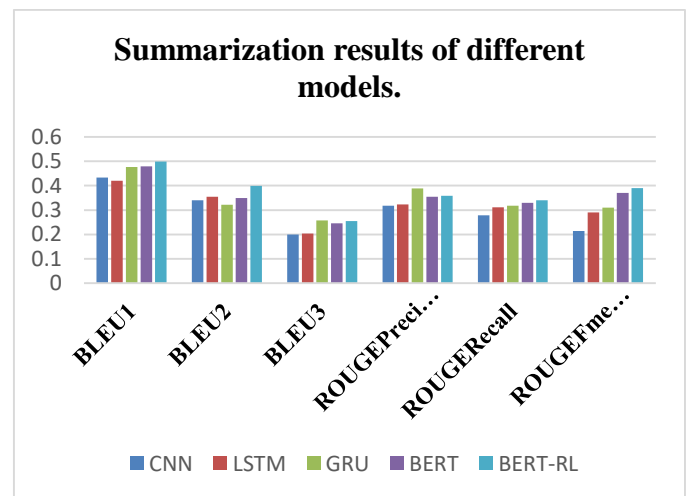

Fig. 4.   Performance comparison experiments measured by BLEU score and ROUGE score.

TABLE II.   MODEL PERFORMANCE WITH VARYING LENGTHS OF SIMULATED CONVERSATIONS

	BLEU1	BLEU2	BLEU3	ROUGE Precision	ROUGE Recall	ROUGEF1- score
BERT-RL(1 Turn)	0.470	0.370	0.320	0.350	0.340	0.354
BERT-RL(3 turn)	0.496	0.389	0.333	0.370	0.351	0.376
BERT-RL(5 turn)	0.520	0.410	0.360	0.380	0.379	0.388
BERT-RL(7 turn)	0.512	0.399	0.350	0.370	0.388	0.89

Table II presents performance metrics for a responsive customer service chatbot system utilizing a hybrid approach of BERT and reinforcement learning (RL) across different numbers of conversation turns. The BLEU scores (BLEU1, BLEU2, BLEU3) measure the similarity between the generated responses and human reference responses, with higher scores indicating better correspondence. Additionally, ROUGE scores (ROUGEPrecision, ROUGERecall, ROUGEFmeasure) evaluate the overlap between the generated and reference responses in terms of precision, recall, and F1 score. As the number of conversations increases, generally, the BLEU and ROUGE scores tend to improve, suggesting that the chatbot's performance benefits from a deeper context and longer interactions. BLEU and ROUGE ratings are critical for evaluating a chatbot's overall performance. Fig. 5 presents BLEU measures the similarity among the chatbot's generated responses and reference responses, specializing in precision. ROUGE emphasizes recall, taking pictures how well the chatbot's responses cover key aspects of the expected solutions. Together, they make sure balanced assessment of accuracy and relevance.
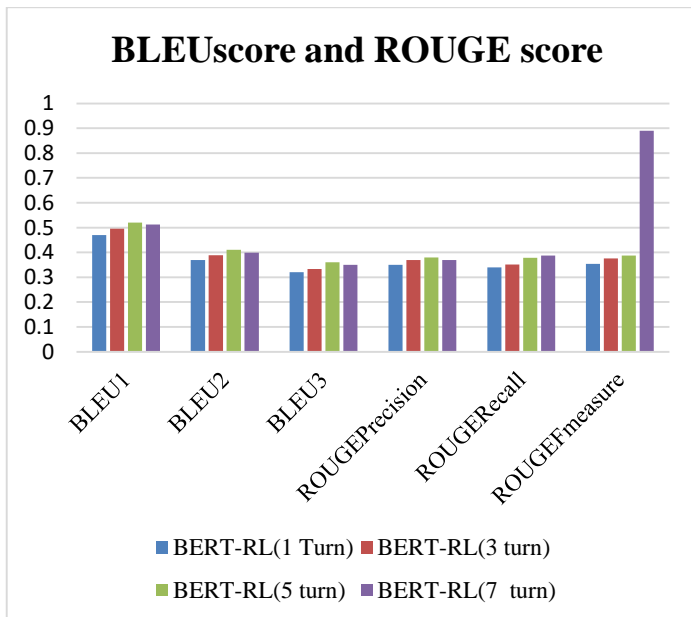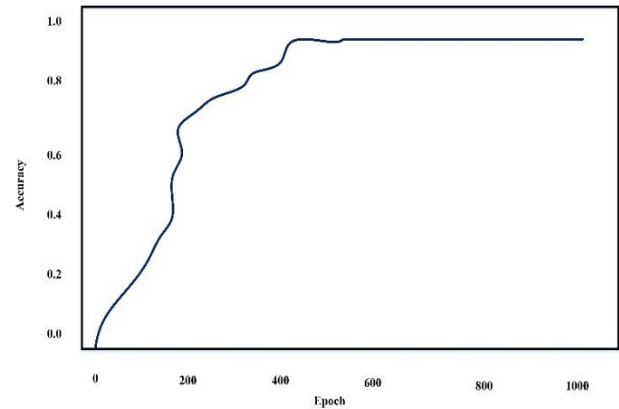

Fig. 6.   Overall accuracy.

The total accuracy of the Customer Service Chatbot built with a BERT and Reinforcement Learning Hybrid System is a comprehensive measure of its ability to understand customer inquiries and respond appropriately this is visually represented in Fig. 6. By combining BERT's sophisticated natural language processing capabilities with reinforcement learning's iterative learning strategy, the chatbot obtains high accuracy in understanding user intent and providing contextual suitable solutions. This hybrid system adapts through user interactions and refines its responses continuously to improve accuracy, resulting in effective and beneficial customer support interactions.

Fig. 5.   Chart displaying performance with various lengths of simulated conversation measured.
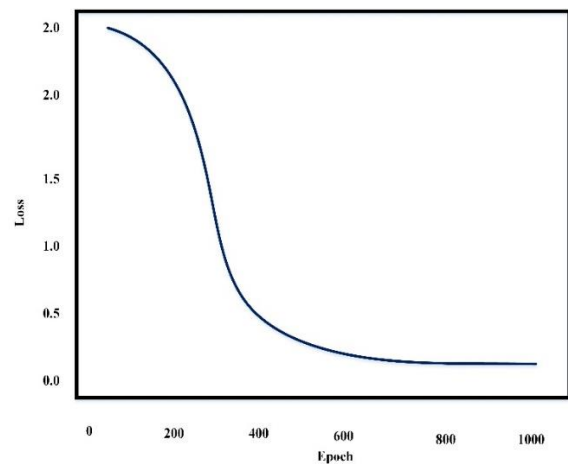

Fig. 7.   Overall accuracy.

The overall loss in the context of a Customer Service Chatbot using a BERT and Reinforcement Learning Hybrid System represents a measure of the discrepancy between the predicted responses generated by the chatbot and the ground truth responses provided by humans is shown in Fig. 7 This loss function quantifies the model's performance during training, guiding it towards minimizing errors and improving response quality. By minimizing the overall loss, the chatbot learns to generate more accurate and contextually relevant responses, thereby enhancing its effectiveness in addressing user queries and improving overall customer satisfaction.

A. Discussion

The conceived work will thus aim at training a Responsive Customer Service Chatbot by applying a hybrid model involving the well-established BERT of language and RL. The idea of this integration is to address a number of drawbacks of the current chatbot systems and improve their outcomes vastly. Most prior chatbots have static decision rules or statistical models, which create limited, contextually superficial conversations that do not extend very well to users individually, or in the shifting trends that correspond to the further course of a conversation. It is also possible with the offer of the proposed chatbot that uses BERT for its functioning since the latter organized language best in paying attention to context and juggling its intricacies. The integration of RL solves the issue of rigidity, which is usually characteristic of normal chatbots [13] [14]. RL enables the chatbot to get smarter by learning from the users' response and feedback thus making it to increase its performance. It is such constant training that assists in improvement of the different and intricate questions that the chatbot is likely to encounter and thus improves the general user experience. Also, the hybrid system is expected to address some issues that may include the overall flow and subject coherence in many turns of the conversation, improved recognition of the user intent, and learning new patterns of the conversation. Doing so benefits from using BERT's deep contextual embeddings combined with RL's adaptive learning mechanisms as more substantive and useful chats are expected to be attained.

VI. CONCLUSION AND FUTURE WORK

The improvement of the Responsive Customer Service Chatbot the use of a hybrid BERT and Reinforcement Learning (RL) machine represents a widespread advancement in chatbot technology. This technique combines the strong language understanding competencies of BERT with the adaptive learning strengths of RL, addressing key obstacles in present day chatbot systems. Traditional chatbots regularly battle with keeping conversational context, knowledge complex person queries, and adapting to dynamic person wishes. By integrating BERT, which excels in processing and decoding nuanced language, with RL, which allows for continuous development based totally on user interactions, the proposed machine aims to deliver a greater clever, responsive, and contextually conscious chatbot. The initial results suggest promising improvements in reaction accuracy and consumer engagement. The BERT model enhances the chatbot's potential to comprehend the intricacies of language, whilst RL best-tunes its responses based on actual-time comments, leading to a greater personalized and effective consumer revel in. This hybrid approach overcomes the stress of rule-based totally structures and the limitations of conventional statistical fashions, imparting a dynamic and scalable solution for customer support applications. Looking forward, future work will consciousness on several key regions. Firstly, expanding the chatbot's competencies to handle greater numerous and complex communication topics could be essential. This entails incorporating additional area-unique know-how and refining the RL algorithms to higher manage problematic user interactions. Secondly, improving the machine's potential to deal with multi-flip conversations and emotional nuances will improve normal person delight. Finally, exploring integration with different superior technology, which includes voice popularity and sentiment evaluation, will similarly raise the chatbot's capability and user experience. Ongoing evaluation and iteration will be critical to ensure the system stays powerful and adaptable to evolving user desires and technological improvements. Future work for the research focus on expand the chatbot's ability to handle a broader range of topics by integrating additional domain-specific knowledge bases and refining the reinforcement learning algorithms for better handling of complex queries. Developing advanced techniques to manage multi-turn conversations, allowing the chatbot to maintain context over extended interactions and respond appropriately to user inquiries that involve emotional nuances.

REFERENCES

[1] S. Ayanouz, B. A. Abdelhakim, and M. Benhmed, "A smart chatbot architecture based NLP and machine learning for health care assistance," in Proceedings of the 3rd international conference on networking, information systems & security, 2020, pp. 1–6.

[2] C. Magoo and M. Singh, "Design and development of new interactive chatbot system for mobile service providers through heuristic-based ensemble learning," Cybernetics and Systems, vol. 55, no. 4, pp. 753–785, 2024.

[3] Z. Wu, Q. She, and C. Zhou, "Intelligent Customer Service System Optimization Based on Artificial Intelligence," Journal of Organizational and End User Computing (JOEUC), vol. 36, no. 1, pp. 1–27, 2024.

[4] M. Demircan, A. Seller, F. Abut, and M. F. Akay, "Developing Turkish sentiment analysis models using machine learning and e-commerce data," International Journal of Cognitive Computing in Engineering, vol. 2, pp. 202–207, 2021.

[5] K. Badran, "Using ChatGPT to Augment Software Engineering Chatbots Datasets".

[6] A.-C. Le, V.-N. Huynh, and others, "Enhancing Conversational Model with Deep Reinforcement Learning and Adversarial Learning," IEEE Access, 2023.

[7] A. Shoufan and S. Alameri, "Natural Language Processing for Dialectical Arabic: A Survey," in Proceedings of the Second Workshop on Arabic Natural Language Processing, Beijing, China: Association for Computational Linguistics, 2015, pp. 36–48. doi: 10.18653/v1/W15-3205.

[8] M. S. Ali, M. W. Anwar, F. Azam, and M. H. Ashraf, "Intelligent Agents in Educational Institutions: AEdBOT–A Chatbot for Administrative Assistance using Deep Learning Hybrid Model Approach," 2024.

[9] A. Al Sallab, H. Hajj, G. Badaro, R. Baly, W. El Hajj, and K. Bashir Shaban, "Deep Learning Models for Sentiment Analysis in Arabic," in Proceedings of the Second Workshop on Arabic Natural Language Processing, Beijing, China: Association for Computational Linguistics, 2015, pp. 9–17. doi: 10.18653/v1/W15-3202.

[10] "Customer Service Chatbot-Data." Accessed: May 11, 2024. [Online]. Available: https://www.kaggle.com/datasets/ngawangchoeda/customer-service-chatbotdata

[11] N. Bhartiya, N. Jangid, S. Jannu, P. Shukla, and R. Chapaneri, "Artificial neural network based university chatbot system," in 2019 IEEE Bombay Section Signature Conference (IBSSC), IEEE, 2019, pp. 1–6.

[12] Q.-D. L. Tran and A.-C. Le, "Exploring bi-directional context for improved chatbot response generation using deep reinforcement learning," Applied Sciences, vol. 13, no. 8, p. 5041, 2023.

[13] Q.-D. L. Tran and A.-C. Le, "Exploring Bi-Directional Context for Improved Chatbot Response Generation Using Deep Reinforcement Learning," Applied Sciences, vol. 13, no. 8, Art. no. 8, Jan. 2023, doi: 10.3390/app13085041.

[14] H. Palivela, "Optimization of paraphrase generation and identification using language models in natural language processing," International Journal of Information Management Data Insights, vol. 1, no. 2, p. 100025, 2021.