Personalized chat-bot Modeling and Performance Analysis: Emulating Specific Character

Lahari Kethinedi
kethinedi.lahari@ufl.edu
University of Florida

Haswanth Popuri
haswanthpopuri@ufl.edu
University of Florida

Chandra Hasan Reddy Busa
chandra.busa@ufl.edu
University of Florida

Harish Thimmapuram
harish.thimmapur@ufl.edu
University of Florida

Shantanu Kapoor
s.kapoor@ufl.edu
University of Florida

Abstract

This paper explores the development of personalized chat-bots that can replicate specific character behaviors, aiming to determine which computational architecture best captures the nuances of character dialogue. We compare two advanced models: a bidirectional LSTM (BiLSTM) with self-attention, known for its ability to process information in both directions and potentially enhance pattern recognition in dialogue, and a decoder-based Transformer model, which is popular in current research for its effectiveness in sequence modeling. Both models are trained on the Harry Potter Dialogue dataset to generate character-specific responses. This specialized training is designed to enhance character-specific dialogue generation, instilling distinct, consistent personalities in chat-bots and addressing common issues like personality inconsistency and generic responses. Our hypothesis is that while the Transformer is anticipated to perform robustly due to its advanced capabilities and current popularity, the BiLSTM might also show strong performance by effectively recognizing bidirectional context patterns. We evaluate these architectures using the BLEU score to assess how well they mimic the targeted character dialogues and to explore their potential in creating engaging and realistic chat-bot interactions. [1]

1 INTRODUCTION

The journey of dialogue systems has been marked by significant evolution, beginning with basic, goal-directed agents and culminating in the intricate, personalized conversational AI we see today. This evolution reflects broader advances within the fields of natural language processing (NLP) and artificial intelligence (AI). Despite these considerable advancements, the interaction between humans and machines remains fundamentally limited, often failing to convincingly simulate human-like exchanges. Notably, while neural models have grown increasingly capable, they typically falter in sustaining meaningful dialogue beyond brief interactions. These limitations, as detailed by (Serban et al., 2015) and (Vinyals and Le, 2015), become apparent as generic models quickly reveal a lack of consistent personality, short-term memory constraints, and a propensity for vague responses. This contrast highlights the ongoing challenges that lie in bridging the gap between the technical capabilities of AI and the dynamic nuances of human conversation.

Addressing these shortcomings, our research pivots towards a nuanced approach by exploring the personalized chat-bot modeling through two pivotal architectures: the BiLSTM model with self-attention and a decoder-based Transformer model. By enhancing the BiLSTM's sequence-to-sequence capabilities with self-attention, we aim to improve the model's focus on pertinent aspects of the input data, thereby enhancing the relevance and coherence of its outputs. Moreover, given the known capabilities of Transformer in generating contextually rich dialogues, our hypothesis centers on the potential elevation in performance through targeted retraining on character-specific data sets, which could instill a more defined personality in chat-bots.

In a study chat-bot (Rakib et al., 2021), designed to deliver empathetic responses, employs a Sequence-to-Sequence (Seq2Seq) encoder-decoder framework where the encoder is based on Bi-directional Long Short Term Memory (BiLSTM). This choice of BiLSTM is driven by its proven capability in previous studies to effectively handle nuanced conversational contexts by processing data in both forward and reverse directions, enhancing the accuracy of the response generation. This superior performance of the

[1] https://github.com/haswanth89/DigitalAvatars-NLP-

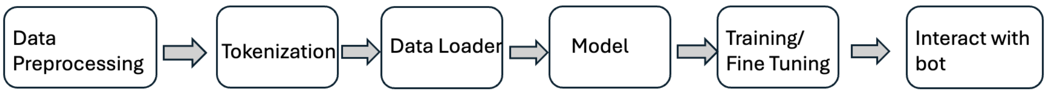

Figure 1: Sequential diagram of the methodology, from data processing to model inference, used in the development of a chat-bot.

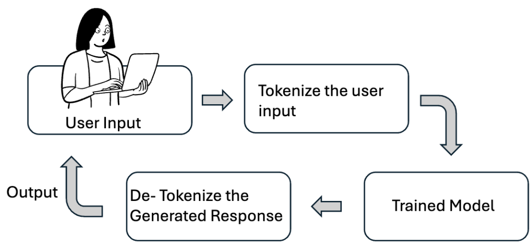

Figure 2: Flow chart depicting chatbot-user interaction, outlining the internal processing of user input and how an output is generated from a trained model.

BiLSTM architecture in generating empathetic interactions underscores its suitability for complex conversational applications.

Our innovative approach involves retraining the Transformer model exclusively with data representing specific characters, diverging from conventional practices of broad dataset training. This method is designed to refine the chat-bot's personality emulation capabilities, thereby addressing common issues identified in prior models such as personality inconsistency and generic responses. The expectation is that both the BiLSTM with self-attention and the modified Transformer will exhibit improved adaptability and specificity in dialogues, reflecting a more realistic and engaging character portrayal.

This paper aims to deliver a comprehensive evaluation of these models, comparing their performance in generating dialogue that not only feels genuine but also remains consistent with the character's defined traits. By conducting extensive experiments and analyzing the outcomes, we seek to contribute meaningful insights into the advancement of chat-bot technology, aiming to bridge the gap between human-like interaction and current machine capabilities in personalized settings.

The succeeding structure of this paper is outlined as follows. Section 2 reviews previous research, discussing relevant developments and identifying existing gaps within the field. Section 3 details the methodologies employed for training both the BiLSTM model with self-attention and the decoder-based Transformer model. Subsequently, Section 4 evaluates the performance of these models using the BLEU score metric. The paper concludes with Section 5, which summarizes our findings and suggests directions for future research.

2 RELATED WORK

The groundwork for dialogue systems was laid by early efforts that focused on executing specific tasks. Highlighted by (Young et al., 2010), these systems utilized dialogue state tracking and response generation to accomplish tasks within narrowly defined domains. The introduction of partially observable Markov decision processes (POMDPs) by Young and colleagues in 2013 (Young et al., 2013) provided a sophisticated framework for enhancing the performance of these goal-oriented dialogue systems. Despite their advancements, these systems primarily neglected the complexity of human-like interactions, such as embodying personality or maintaining long-term engagement, focusing instead on executing tasks like booking flights or arranging appointments.

For the chit-chat setting, the most relevant work is (Li et al., 2016). For each user in the Twitter corpus, personas were captured via distributed embeddings (one per speaker) to encapsulate individual characteristics such as background information and speaking style, and they then showed using those vectors improved the output of their seq2seq model for the same speaker.

The transition towards personalization in dialogue systems gained momentum, driven by the need to adapt responses to the unique characteristics of users, particularly in informal conversation contexts. The advent of the Transformer model and its variation, TransferTransfo, introduced by (Wolf et al., 2019), exemplified this shift. Pre-trained on extensive corpora and fine-tuned for personalization, these models demonstrated enhanced capabilities in generating engaging and coherent dialogues. The creation of the Persona-Chat dataset by (Zhang et al., 2018) and subsequent research efforts further propelled the development of agents that could tailor their speaking styles and content based on explicit user profiles, fostering more personalized interactions.

Additionally, innovations in variational autoencoders (VAEs) and GPT-n architectures have facilitated more sophisticated and dynamic personalization (Connor et al., 2021). These advancements enable the generation of responses that not only reflect the user's interests and personality but also vary in tone and style, making conversations feel more natural.

Moreover, integrating reinforcement learning (RL) with dialogue personalization has unveiled new possibilities for adjusting conversational strategies based on metrics of user engagement and satisfaction (Zhao et al., 2019). RL-based methods focus on optimizing long-term interaction outcomes through learning from user feedback, allowing dialogue systems to align more closely with individual user preferences over time.

However, the transformer's heavy reliance on extensive parameterization and large training datasets poses challenges for deployment in real-time environments equipped with low-resource hardware. Addressing this issue may involve revisiting the BiLSTM architecture with an attention mechanism, which offers potential solutions to these memory and parameter-related challenges. The BiLSTM model typically does not demand as much data for training, particularly for context-specific applications, making it a more manageable option. Fine-tuning a Transformer model with billions of parameters can be cumbersome, requiring significant computational resources for updates. Hence, exploring the BiLSTM approach presents a viable avenue to mitigate these concerns.

	Model 1	Model 2
Algorithm	Deep Neural Network	Transformer
Main Technique	Sequence to Sequence (Seq2seq)	Casual Language Modeling(LM)
Enhancement Techniques	Bidirectional LSTM with attention Mechanism	AutoModelForCausalLM

Table 1: Algorithm and technique used in this experiments.

3 METHODOLOGY

In this paper, we delineate the comprehensive methodology employed for the development of a sophisticated conversational agent, herein referred to as a chat-bot. The process initiates with an extensive data collection phase, where diverse textual datasets are compiled, encompassing chat logs, conversation transcripts, and other pertinent text data across about a specific character. This is followed by a meticulous data preprocessing stage designed to refine the quality of the input data. Here, textual data undergo cleaning processes to remove extraneous characters and symbols, normalization through lower-casing, and optional steps such as stop-words removal and spell correction to enhance data homogeneity.

Subsequent to preprocessing, the tokenization phase involves the conversion of cleaned text into a structured format suitable for model training. Depending on the desired granularity, tokenization may be implemented at the word, subword, or character level. The tokenized data is then organized into batches using a data loader, which also ensures random shuffling to mitigate bias during model training.

The core of the chat-bot framework is its model architecture, in this study are based on neural networks such as Bidirectional Long Short-Term Memory (BiLSTM) network or Transformer.

This architecture incorporates multiple layers and connections, including advanced mechanisms like attention, all initialized with pre-defined parameters and optimized through training. Model training constitutes a critical phase where the chat-bot is exposed to the data repeatedly in multiple epochs to minimize the predefined loss function using optimizers such as Adam or SGD. Throughout this phase, performance is intermittently assessed on a validation set to monitor and guide the training process effectively.

Upon completion of the training phase, the chat-bot is equipped to interact with users. This interaction is facilitated through a user interface that collects input, which the model processes using its trained weights to generate and return conversational responses. Each component from data handling to response generation is crucial to the functionality and efficacy of the chat-bot, illustrating a complex but systematic approach in the design and development of AI-driven conversational agents.

Our methodical framework delineates the progression from preliminary data management to the deployment of an interactive chat-bot, as depicted in Figure 1. Each essential stage of the development process is meticulously traced, offering a clear visualization of the workflow. Additionally, the mechanics of chat-bot interaction are elucidated in Figure 2, providing a comprehensive understanding of the operational dynamics involved in user engagement.

3.1 BIDIRECTIONAL LSTM CHAT-BOT

The Bidirectional LSTM chat-bot is a kind of the conversational agent or the chat-bot whose Bidirectional LSTM neural network architecture is used for sequential data processing centering on a natural language. Unlike classical LSTM architectures, which are unidirectional and accept a sequence input in a forward just path (from past to future), Bidirectional LSTM accepts sequences in both directions (from past to future and from future to past) simultaneously. This means that the model makes processing to become bidirectional and be able to get the context dependency from both preceding and consecutive words of the input sentence.

LSTM chat-bots conversely use the recurrent neural network (RNN) to have bidirectional connections that exploit contextual dependencies from both before reaching the current the token to the forward token in the input sequence. The architecture foundations comprise of 2-way LSTM layers preceded by the embedding layer, carried out the vectorized tokenization. An attention mechanism, on which the focus could be on the relevant parts of the input sequences, is an optional component to increase the model's performance. Bidirectional LSTMs work for processing sequences linearly. This is a benefit due to the need to capture dependencies and sequential inner structure which are predominant in the natural language data.

Bi-directional LSTMs process sequences both forward and backward in time. That is why they use temporal and recurring connections to store and pass contextual information to others. As an upshot, backpropagation through time (BPTT) is implemented in the training stage to accelerate the learning process with looped gradients computed of timeline-way. Sequence generation on Box Forwards LSTM chat-bots is accomplished predicting of the next token in the sequence based on preceding context and target outcome.

3.2 TRANSFORMER CHAT-BOT

Chat-bot which consists of transformer architecture and is one of the applications of chat-bot for conversations is determined as a transformer chat-bot. The Transformer architecture, introduced by the paper "Attention Mechanisms All you Need" by (Vaswani et al., 2023) prevailing recurrent connections and replacing them with self-attention mechanisms.

The self-attention mechanism adopted Transformer architecture to challenge context connection by eradicating recurrent links. It features two components: an encoder and a decoder. The encoder and the decoder subsystems have a multilayer structure comprising the self-attention mechanism and the feedforward neural network modules, respectively. Since the attentions of different heads in muti-head self-attention mechanisms are processed in parallel and contexts are aggregatively modeled by the Transformer architecture of each layer, comprehensive contexts can be modeled in it.

Transformers process input tokens in paral-

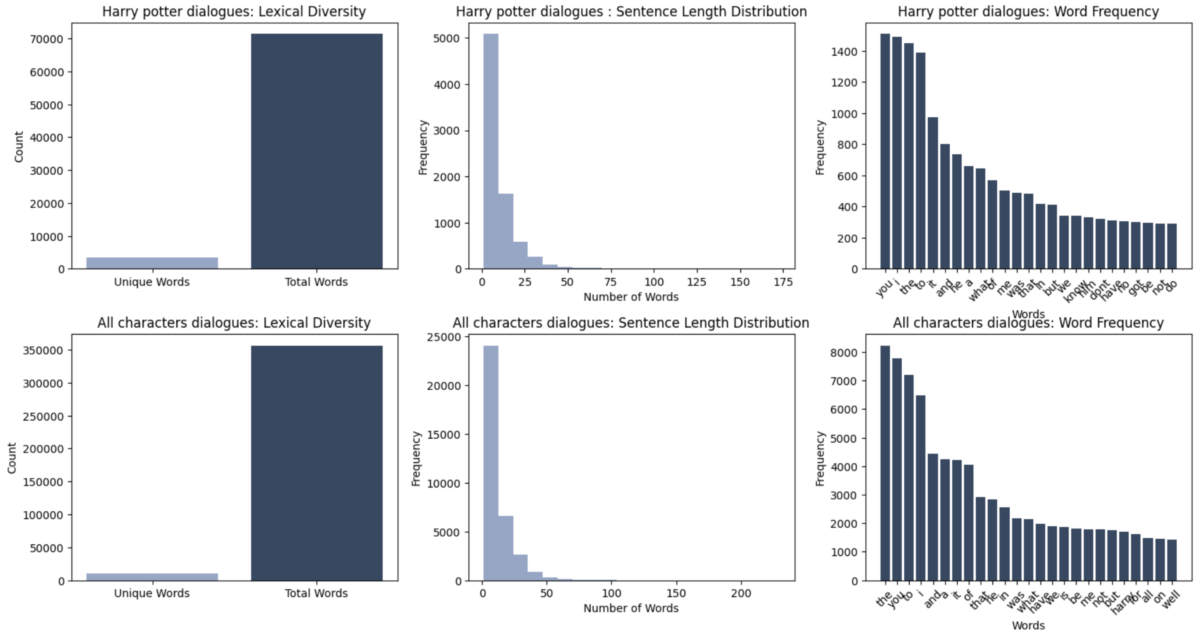

Figure 3: The image is a set of six histograms analyzing the dialogues from the "Harry Potter" series. The top three charts represent Harry Potter's dialogue, showing his lexical diversity with 3,446 unique words out of 71,360 total words, the distribution of sentence lengths favoring shorter sentences, and word frequency highlighting common English words. The bottom three charts represent all characters, revealing greater lexical diversity with 10,857 unique words out of 356,657 total words, a similar preference for shorter sentences, and word frequency with common words.

lel instead of sequentially which highlights the right manner of processing words at a whole sentence level with long-term relationships between them. The application of automated techniques, for example, masked language modelling where it is required to guess the consecutive tokens in target sequences while keeping the subtle of the sentence is also exercised during the training. Positional encoding vectors are assembled with token embeddings to provide positional cues essential for seeking linear processing in language chatting based on the Transformer model.

4 EXPERIMENTS

4.1 DATASET

The dataset utilized for this study was sourced from the Harry Potter Dialogue (HPD)[2], which includes all dialogue sessions from the Harry Potter novels in English. This dataset comprises 1,042 training dialogue sessions, each with one positive response, and 149 testing sessions. Each session is annotated with detailed background information such as dialogue scenes, speakers, character relationships, and attributes, providing a dynamic understanding of each character's development throughout the series.

The dataset from the Harry Potter series contains 354 distinct characters and includes 15,819 conversational exchanges. For an in-depth evaluation of lexical diversity and various other characteristics and parameters of the dataset, such as comparisons of dialogues involving Harry versus all other characters, refer to Figure 3.

To prepare the HPD dataset for chat-bot modeling, the following preprocessing steps were implemented: The data underwent a cleaning process to remove special characters, formatting irregularities, and redundant elements. Lastly, the cleaned text was tokenized using subword tokenization to more effectively manage the unique vocabulary, including magic spells and character names inherent to the Harry Potter series.

4.2 BIDIRECTIONAL LSTM CHAT-BOT

In this study, we employed a BiLSTM architecture augmented with self-attention, specifically designed to handle dialogue-based text from the Harry Potter dataset. The model leverages GloVe embeddings pre-trained on Wikipedia large dataset, which provide a rich semantic

[2]https://nuochenpku.github.io/HPD.github.io/

base and aid in dimensionality reduction, thus facilitating faster convergence during training. The model's core comprises 3 LSTM layers (2 encoder and 1 decoder), arranged to process sequences both forwards and backwards, integrated with a self-attention layer that emphasizes relevant contextual cues. Outputs are generated through a dense layer with softmax activation.

The architecture of our BiLSTM network consists of multiple LSTM layers; specifically, a stack of two Bidirectional LSTM layers, each with 50 units. These layers are adeptly configured to process text sequences bidirectionally, ensuring comprehensive learning from both preceding and subsequent contexts within the text. Interspersed with these layers are Dropout layers with a rate of 0.5 to mitigate the risk of over-fitting by randomly omitting a portion of the neural units during the training phase, thereby forcing the model to learn more robust features.The model's hyperparameters are carefully chosen to optimize its learning capability and performance. The embedding layer maps each word to a 100-dimensional vector, striking a balance between capturing semantic details and computational efficiency. The Bidirectional LSTM layers, with their 50 units, are strategically sized to discern complex patterns within the dialogue without becoming computationally prohibitive.

Hyperparameter tuning was pivotal to our model's efficacy. We utilized a small batch size of 8, which fortified the stability of gradient estimates and thereby bolstered the model's generalization capabilities. The training spanned over 300 epochs, featuring an initial learning rate of 0.001, in conjunction with the Adam optimizer, known for its efficiency with sparse gradients. A learning rate scheduler reduces the rate if no improvement in validation loss is observed over three epochs, ensuring efficient convergence. Training utilizes cross-entropy loss, suitable for the probabilistic outputs of our model.

4.3 TRANSFORMER CHAT-BOT

In our study, we fine-tuned a decoder based pre-trained transformer model from the AutoModelForCausalLM hugging face library, renowned for its robust language generation capabilities, to serve as a dialogue-based chat-bot using the Harry Potter dataset. We leveraged the Harry Potter dataset to retrain the transformer model, applying a specific set of hyperparameters and training strategies to refine its performance and make it suitable for engaging in human-like conversations.

We started with the standard decoder Transformer architecture and retrained it using a learning rate of 2e-5, suitable for fine-tuning to make subtle yet effective updates. We employed the Adam optimizer alongside a learning rate scheduler that reduced the rate as validation loss plateaued, using categorical cross-entropy as the loss function for next-word prediction. Training proceeded with early stopping based on performance on the development set to avoid over-fitting.

In the development of our chat-bot, we meticulously set several hyperparameters to enhance the dialogues' naturalness and relevance. The '$no\_repeat\_ngram\_size$' was configured to 3 to prevent repetitive word sequences, ensuring varied and engaging conversations. We also controlled the randomness and creativity of the responses using the 'top_k' and 'top_p' settings, limiting word choices to the most probable 100 and refining potential responses to those surpassing a 70% cumulative probability threshold, respectively. Additionally, the 'temperature' setting of 0.8 moderated the probability distribution, favoring more probable words to maintain coherence without sacrificing unpredictability. Together, these parameters were optimized to produce concise, relevant, and lifelike interactions, capped at a maximum length of 1000 tokens to avoid overly verbose outputs, making the chat-bot appear more intuitive and engaging. These settings help in mimicking a more natural conversation flow, making the chat-bot seem more lifelike and responsive.

4.4 EXPERIMENTAL RESULTS

The Transformer model exhibited superior results, achieving a BLEU score of approximately 0.58. This outcome indicates a moderate capability of the chat-bot to produce syntactically relevant dialogue but also demonstrates narrative voice of Harry Potter. In comparison, the Self-attention based BiLSTM model scored a 0.2428, which, while indicative of some alignment with the expected dialogue, suggests a less congruent replication of the character's language style.

The stark contrast in performance—where the Transformer's score was more than twice that of the BiLSTM model—can be attributed to the Transformer's utilization of a large pre-trained language corpus. This extensive background enables the model to generate and fine-tune dialogue more effectively. Meanwhile, the BiLSTM model, despite employing GLoVE embeddings to capture word meanings, struggled with long-range dependencies essential for capturing the intricate language patterns of the Harry Potter dialogue.

The challenge becomes evident when examining the output generated by these models. While there is a degree of stylistic resemblance, the responses often appear out of context, revealing a critical hurdle in dialogue generation. To achieve linguistic coherence, models need an architecture that accounts for a wide array of situational contexts and environmental conditions that the characters might encounter such as engaging in a duel, attending class or expedition. This ensures responses that are not only syntactically correct but also semantically meaningful.

The results underscore the Transformer model's advanced capabilities in language modeling and hint at its suitability for applications demanding high linguistic fidelity, such as personalized chat-bots. The intent behind employing these models was ambitious: to move beyond repeating famous lines and toward crafting novel responses that capture the character's unique style while remaining pertinent to user inputs. Realizing this goal necessitates sophisticated techniques like fine-tuning on extensive datasets and integrating external knowledge for contextual depth. For these sophisticated techniques the transformer models can act as a starting point of comparison.

Model	BLEU Score
BiLSTM	0.2428
Transformer	0.5827

Table 2: BLEU Scores of BiLSTM and Transformer Model

5 CONCLUSION

In conclusion, our comparative analysis of training dialogue chat-bots using the Harry Potter dataset demonstrated that the transformer-based model outperformed the BiLSTM model with self-attention. The transformer's advantage stems largely from its extensive pre-training on a diverse corpus, which provided a deep foundational knowledge of language, enhancing its ability to dynamically adapt during fine-tuning to the specific nuances of the Harry Potter dialogue. In contrast, the BiLSTM model, despite effectively handling sequential data and being enhanced with GloVe embeddings for a basic linguistic understanding, could not match the adaptability and depth of language comprehension offered by transformer. This disparity highlights the significant impact of pre-training on model performance in complex dialogue generation tasks.

A comprehensive analysis of the performance of

Character	Trained Model Dialogue Responses
User:	We're going to the park
Harry:	Here he comes, the birthday boy.
User:	What are goblins
Harry:	Clever as they come goblins but not the most friendly of beasts.
User:	This is what we're going to do...
Harry:	is that when we go out, we're gonna buy you two new presents.
User:	Can you hear me?
Harry:	It's just, I've never talked to a snake before.

Table 3: Output Sequence of Transformer Model. User indicates user text and Harry indicate the reply we get from the trained model.

Character	Trained Model Dialogue Responses
User:	make Harry get it.
Harry:	Come on.
User:	I want to read that letter
Harry:	Yeah, sure
User:	Good afternoon
Harry:	Welcome to your first flying lesson.
User:	Get out of the way,
Harry:	Move
User:	I can't see anything.
Harry:	!!!,!!,!!?!!

Table 4: Output Sequence of BiLSTM Model. User indicates user text and Harry indicate the reply we get from the trained model.

both models is presented in Table 3 and Table 4. These tables illustrate the quality of human interaction with the trained weights of each model, providing insight into their respective capabilities in simulating human-like conversations.

Training a model for machine translation as a proxy for dialogue generation presents numerous challenges, not least of which is the extensive computational resources and sophisticated hardware required. Additionally, optimizing the model's hyperparameters is a significant hurdle that can impact overall performance.

In this study, the development of a generic chatbot is explored through the lens of machine translation. This approach does not account for the histories of previous interactions, which could limit the model's effectiveness in longer conversations. Frequently, the outputs generated were found to be repetitive and lacked specificity. Furthermore, the absence of high-quality, real-life conversational data further compromised the chatbot's ability to emulate human-like interactions. Many responses were also excluded from analysis due to their excessive length or inconsistencies, indicating potential areas for model refinement.

6 FUTURE WORK

In future developments of our dialogue chat-bot project, we aim to carefully select and refine our datasets to better match the model's requirements and computational capacities. This will involve a thorough analysis of data quality and relevance prior to modeling, ensuring that we can streamline the training process and make it more efficient.

We also plan to explore sophisticated embedding techniques that incorporate scenario situations, location details, and personality traits. These enhancements will enable the chat-bot to deliver responses that are not only coherent but richly contextualized within the conversational setting. This approach will significantly improve the chat-bot's adaptability and interaction quality, opening up new avenues for personalized communication in conversational AI.

To enhance the performance of chatbots, the implementation of advanced mechanisms, such as the Luong attention mechanism, is advocated. This method has been supported by various scholarly articles. Experimentation with diverse hyperparameters and evaluation metrics might also yield improvements in chatbot functionality.

7 ACKNOWLEDGEMENTS

We extend our sincere appreciation to both the leadership and mentor for their invaluable guidance and support throughout this research endeavor. To the leadership team, including Dr.Bonnie J Dorr, Professor Sangpil Youm, Chathuri Jayaweera and Justin Ho, we express our gratitude for their dedicated mentorship and unwavering support, which have played a pivotal role in shaping our methodology and ensuring the rigor of our research. Similarly, we are thankful to the mentor Venkat for their wisdom and encouragement, which have been instrumental in our growth and development throughout this journey. Their commitment to excellence and willingness to share expertise have been integral to our progress, and we deeply appreciate their contributions.We also acknowledge the University of Florida for its institutional support, providing access to facilities and resources that have facilitated our research endeavors.

References

Marissa Connor, Gregory Canal, and Christopher Rozell. 2021. Variational autoencoder with learned latent structure. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pages 2359–2367. PMLR.

Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2016. A persona-based neural conversation model. *CoRR*, abs/1603.06155.

Afsana Binte Rakib, Esika Arifin Rumky, Ananna J. Ashraf, Md. Monsur Hillas, and Muhammad Arifur Rahman. 2021. Mental healthcare chatbot using sequence-to-sequence learning and bilstm. In *Brain Informatics*, pages 378–387, Cham. Springer International Publishing.

Iulian Vlad Serban, Alessandro Sordoni, Yoshua Bengio, Aaron C. Courville, and Joelle Pineau. 2015. Hierarchical neural network generative models for movie dialogues. *CoRR*, abs/1507.04808.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2023. Attention is all you need.

Oriol Vinyals and Quoc V. Le. 2015. A neural conversational model. *CoRR*, abs/1506.05869.

Thomas Wolf, Victor Sanh, Julien Chaumond, and Clement Delangue. 2019. Transfertransfo: A transfer learning approach for neural network based conversational agents. *CoRR*, abs/1901.08149.

Steve Young, Milica Gašić, Simon Keizer, François Mairesse, Jost Schatzmann, Blaise Thomson, and Kai Yu. 2010. The hidden information state model: A practical framework for pomdp-based spoken dialogue management. *Computer Speech Language*, 24(2):150–174.

Steve Young, Milica Gašić, Blaise Thomson, and Jason D. Williams. 2013. Pomdp-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, 101(5):1160–1179.

Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. Personalizing dialogue agents: I have a dog, do you have pets too? *CoRR*, abs/1801.07243.

Tiancheng Zhao, Kaige Xie, and Maxine Eskénazi. 2019. Rethinking action spaces for reinforcement learning in end-to-end dialog agents with latent variable models. *CoRR*, abs/1902.08858.