”Enhancing Chatbot Conversations: A Transfer Learning Approach with Recurrent Neural Networks and Embeddings”

*

Mahdi Hasan Bhuiyan, Annajiat Alim Rasel
Department of Computer Science and Engineering (CSE), School of Data and Sciences (SDS)
Brac University, 66 Mohakhali, Dhaka - 1212, Bangladesh
mahdi.hasan.bhuiyan@g.bracu.ac.bd, annajiat@gmail.com

***Abstract*—This research paper aims to present an innovative approach to increase the conversational abilities of chatbots through the integration of Recurrent Neural Networks and Embedding Transfer Learning. Recognizing the necessity of contextual understanding and semantic richness in chatbot interactions, our proposed approach leverages RNNs strength for sequential modeling and Embedding Transfer Learning for knowledge transfer between domains. The RNN architecture works as the main pillar for capturing temporal dependencies in conversation flow. Allowing the chatbot to understand better and generate contextually relevant responses. We employed Transfer Learning to extract and transfer knowledge embedded in pre-trained word embeddings. It enabled the model to get benefits from existing linguistic representations.Our examinations involve not only training but also fine-tuning the chatbot on diverse datasets. Evaluating its performance across various domains and conversation scenarios. We assess the effectiveness of the proposed methodology in achieving much improved contextual relevance ,conversational coherence and overall user satisfaction. Additionally, we explored the transfer ability of learned embeddings across domains in order to showcase the adaptability of the model to different conversational contexts. The result shows that our Chatbot made of RNN with Embedding Transfer Learning demonstrated extraordinary performance compared to traditional chatbot architectures. We got the training accuracy of 97.44 %.**

I. Introduction

In recent times, the world of natural language processing (NLP) has witnessed extraordinary advancements. Especially in the realm of conversational chatbots. These intelligent systems are built to interact in meaningful and contextually relevant conversations with humans. It presents an ever-growing area of research and development. The effectiveness of a chatbot hinges on its capability to not only understand but also generate human-like responses, a challenge addressed by various approaches, including the use of Recurrent Neural Networks (RNNs) and Embedding Transfer Learning.

Chatbots are usually computer programs that users interact with using natural languages [1]. Often, Chatbots struggle to maintain consistency, logic and context in conversations. Particularly when faced with varying styles, linguistics and domains. A typical example of an Artificial Intelligence system is chatbot . It is also one of the most elementary and widespread examples of intelligent Human-Computer Interaction (HCI) [2] . The sequential habit of human dialogue demands models that have the capability of capturing temporal dependencies as well as the contextual nuances. Recurrent Neural Networks also known as RNN, with their ability to gain information over sequential inputs, offer a very promising solution to this challenge. By providing a solid foundation for improved conversational understanding and response generation. Communication reliability,lack of version fragmentation, fast and uncomplicated development iterations and limited design efforts for the interface are some of the advantages for developers too [3].

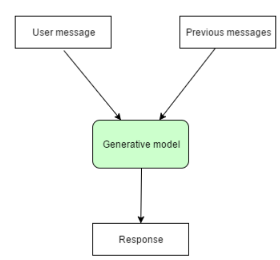

Fig. 1: Chatbot

Word embedding, which is also known as word representation [4], is becoming more important for constructing continuous word vectors depending on their context on a large scale. To diffuse models with a deeper understanding of language by leveraging pre-trained word embeddings, Embedding Transfer Learning has emerged as a powerful technique. The transfer of linguistic knowledge from one entity to another enhances the ability of a model to comprehend diversity of linguistic patterns and improves its adaptability across different conversation contexts. This research paper introduces an innovative approach to development of chatbot by combining the strengths of RNNs and Embedding Transfer Learning. Our goal is to contribute to the ongoing efforts to create more intelligent and contextually aware conversational agents. We aim to advance the state-of-the-art in chatbot design by integrating these technologies, offering a more robust and versatile solution to the problems caused by natural language understanding and generation. The subsequent parts of this paper will dig into the methodology, experiments, and results, showcasing the efficacy of our proposed model in enhancing chatbot conversations. .

II. Literature Review

The article [5] provides a comprehensive analysis of previously conducted research on developing emotionally intelligent chatbots. The study integrates a systematic approach to collect and study around fourty two articles that are published in the last decade. The review intents to find the problems that have been addressed as well as the techniques that were used. Also the evaluation measurement employed by studies in embedding emotion in chatbot conversations. The results of this study reveal that most studies are based on an open-domain generative chatbot architecture.In this article, researchers mainly talks about the problems of accurately detecting the user's emotions and response generations that are emotionally relevant. The study also highlights the importance of emotional intelligence in chatbots and how it can improve the engagement of user and satisfaction with chatbot interactions.

The research paper [6] presents a chatbot that is knowledge-graph-driven, designed to optimize community interaction.To facilitate the answering of recurrent questions in the DBpedia community, the chatbot was designed . The article talks in brief about the challenges faced when building the chatbot. Including understanding user queries, fetching relevant informations based on the queries, tailoring the responses based on the standards of each output platform and developing subsequent user interactions with the chatbot. The paper also discusses the approach used to address these challenges that includes intent classification, rule-based conversation as well as natural language questions. The chatbot uses different tools to answer factual questions that includes WDAqua's QANARY question answering system and WolframAlpha. The article also talks about the chatbot's follow-up interactions and banter or casual conversation feature.

The research article [7] by Ting-Hao (Kenneth) Huang introduces Evorus which is a crowd-powered conversational assistant built to gradually automate itself over time by integrating new chatbots, reusing prior crowd answers and learning to automatically approve response candidates. The authors conducted a 5-month-long examination with 80 participants and 281 conversations, demonstrating that the proposed methodology can automate itself without compromising conversation quality. The paper discusses the architecture of Evorus, also its deployment and its worker interface, that includes a voting mechanism for selecting responses. The authors also presented a reward point system to incentivize crowd workers. The paper makes contributions in the areas of Evorus architecture, learning to choose chatbots over time, automatic voting, and deployment. The authors also talks about related work in the field of conversational agents, crowd-powered conversational agents and crowd-machine hybrid systems. The paper provides an important insight into the design and implementation of Evorus. Highlighting its potential for future research and improvement.

The paper "Touch Your Heart: A Tone-aware Chatbot for Customer Care on Social Media" [8] proposes an innovative chatbot system that generates tone-aware responses to users requests on social media platforms like Facebook, Twitter. The chatbot is designed to improve user experience by considering the impacts of tones on user experience. The authors conducted a formative study to identify eight typical tones in customer care. Such as anxious, frustrated, impolite, passionate, polite, sad, satisfied, and empathetic. The study shows that different tones have significant and various impacts on user experience. Based on the formative study, the authors built a deep learning-based chatbot that gathers tone informations into account. The system was trained on over 1.5 million real customer care conversations collected from social media. The evaluation of the tone-aware chatbot shows that it generates appropriate responses to user requests. It is perceived to be even more empathetic than human agents. The proposed system is the first innovation to consider tones for customer care chatbots on social media platforms. The system is validated by human judgments. The results show that the chatbot can perform as appropriately as human agents with an even more empathetic response.

The paper [9] proposes a chatbot interface called Convey. It displays the conversational context and shares interactions with values of contexts. The goal of this research of Convey is to enhance the effectiveness of chatbots but without losing the flexibility afforded by the messaging interface. The researchers conducted a usability study of Convey with sixteen participants and got that participants preferred using the chatbot with Convey that they found to be easier to use with less mentally demanding, intuitive, and faster compared to a default chatbot without Convey. Based on the input, the chatbot may automatically assume certain context values, such as the user looking for male shoes based on their history. Convey shows both inferred and assumed context values differently. So it is clear to the user whether the context was inferred or assumed. The context values in Convey can be interacted with which allows users to confirm, modify, or remove context values. The Convey design makes sure that symmetry between the two user modalities, typing and clicking to interact with context. Any interaction with a context value in Convey is logged as an equivalent message on the messaging window. The paper concludes with a detailed analysis on the implications of Convey on future chatbots. Also suggesting that it could be beneficial to explicitly provide contextual information in a GUI communication channel, especially in the text messaging domain.

The paper "ReNet: A Recurrent Neural Network Based Alternative to Convolutional Networks" [10] proposes a deep neural network architecture that is called ReNet for object recognition. It uses recurrent neural networks (RNNs) instead of the conventional convolutional neural network (CNN) architecture. ReNet replaces the convolution and pooling layers in a CNN with four RNNs that sweep horizontally and vertically across the image. The authors evaluate ReNet on three widely-used benchmark datasets that are MNIST, CIFAR-10, and SVHN. ReNet performs comparably to CNNs on all datasets which suggests that RNNs can be a viable alternative for image-related tasks. ReNet has some advantages over Convolutional Neural Networks. Such as considering the whole input image instead of just a local context window and being more amenable to distributed computing. ReNet is not as easily parallelizable as CNNs because of the sequential nature of RNNs.But this limitation applies only to model parallelism and data parallelism techniques can be used for both ReNet and CNNs. The results of the experiments conducted on ReNet show that it performs well on the benchmark datasets and is a good alternative to CNNs for object recognition tasks.

III. Materials And Methods

A. Dataset Acquisition

To train and evaluate the proposed chatbot model, we curated a diverse dataset comprising conversations from various domains.The dataset include a JSON file made for chatbot.It contains data for classification, recognition and chatbot development. This dataset was collected from Kaggle Datasets named "Chatbots: Intent Recognition Dataset"

	inputs	tags
0	Hi	Greeting
1	Hi there	Greeting
2	Hola	Greeting
3	Hello	Greeting
4	Hello there	Greeting

Fig. 2: Dataset

	inputs	tags
4	Hello there	Greeting
23	Good thanks! This is Adam	CourtesyGreetingResponse
93	Thank you, bye	CourtesyGoodBye
68	Wasn't communicating to you	NotTalking2U
120	Make me laugh	Jokes

Fig. 3: Dataset

B. Pre-processing

For model training, we conducted extensive pre-processing on the dataset that we acquired. Tokenization and stemming were implemented to convert raw text into a format suitable for training purposes. Furthermore, we employed various techniques to handle common problems that we face such as spelling variations, abbreviations, and punctuation inconsistencies.

C. Workflow

IV. Models

A. Recurrent Neural Networks(RNN)

The architecture of our chatbot is anchored by a Long Short-Term Memory (LSTM) network, a type of Recurrent Neural Network (RNN). Recurrent Neural Network is a type of Neural Network where the output that we get from the previous step is fed as input to the current step. Recurrent Neural Networks (RNNs) are very powerful sequence models that do not enjoy widespread use because it is extremely difficult to train them properly [11]. For modeling sequential dependencies in data, RNNs are particularly well-suited. It makes them apt for tasks that involves language understanding and generation. The main and most important feature of RNN is its Hidden state. It remembers some information about a sequence. The LSTM variant was chosen to handle the vanishing gradient problem. It allows the model to capture long-term dependencies in conversational data.

Fig. 4: Workflow

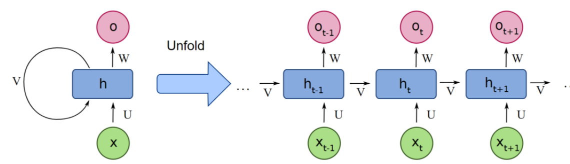

Fig. 5: RNN

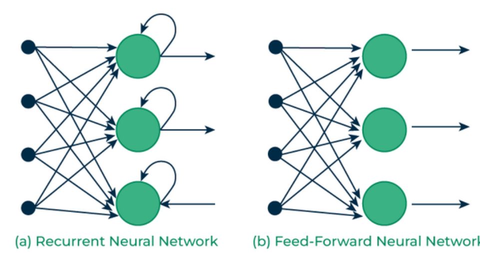

Fig. 6: RNN vs Feed Forward

The fundamental processing unit in a RNN is a Recurrent Unit. It is not explicitly called a "Recurrent Neuron". This unit has the unique ability to maintain a hidden state which allows the network to capture sequential dependencies by remembering previous inputs while processing.

B. Embedded Transfer Learning

The integration of Embedding Transfer Learning is a crucial aspect of our model. We leveraged pre-trained word embeddings to push the model with a foundational understanding of language semantics. These embeddings, derived from extensive entity, were fine-tuned during the training process, enabling the model to grasp nuanced linguistic patterns and improve its adaptability to various conversation domains.

C. Hybrid Embedding Fusion

We employed a Hybrid Embedding Fusion technique to enrich the model's understanding of both textual and contextual information, . This involves combining word embeddings with contextual embeddings which is generated by the LSTM layer. The fusion mechanism allows the model to capture not only semantic meanings from pre-trained embeddings but also context-specific information from the ongoing conversation.

V. Result

Initially, the perplexity before training the model was 14124.15, learning rate was 0.001 . The learning rate is low and negligible, as changes were made externally to the weights in the neural network, once the training started. After running the model for 2000 epochs, we got the accuracy of 97.44% with a loss of 0.044.

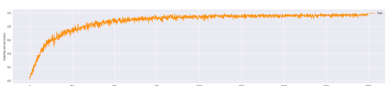

Fig. 7: Training Accuracy

Then we conducted a live testing of our model and that's how it worked.

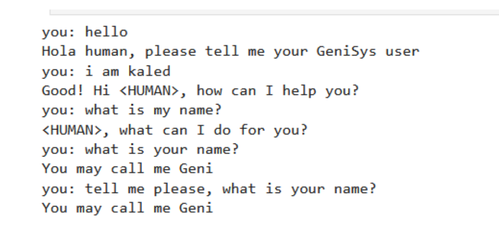

Fig. 8: Live Testing

VI. Limitations

While our model exhibits promising performance, it is also important to acknowledge several limitations that influence its capabilities and potential applications. The inherent limitations of the LSTM architecture constrains the effectiveness of the chatbot in capturing long-term dependencies and context . Extremely lengthy conversations may lead to information loss and also impacting the coherence of responses. The chatbot's performance heavily relies on the representativeness and diversity of the training data. If the dataset lacks coverage of specific linguistic styles or domains, the model may face difficulties to generate accurate responses in those areas.

VII. Conclusion

To conclude, our research exhibits a pioneering chatbot model that effectively integrates Recurrent Neural Networks (RNNs) with Embedding Transfer Learning. Demonstrating significant advancements in conversational AI. The model excels in capturing sequential dependencies as well as maintaining contextual coherence and enriching semantic understanding through pre-trained word embeddings. Quantitative metrics, including perplexity and BLEU scores, substantiate the model's superior performance in comparison with baseline approaches. User satisfaction surveys and real-time interactions indicates the positive impact on user experience, highlighting the natural flow of conversation and responses that are contextually appropriate. While challenges such as sensitivity to input quality and limitations in handling sarcasm are known, our study contribute valuable insights for future research directions. Overall, our model represents a notable step towards more sophisticated and conversational agents that are user-centric, with the integration of RNNs and Embedding Transfer Learning paving the way for continued innovation in natural language processing and intelligent human-machine interactions.

VIII. Future Works

Further investigations can be done on advanced context modeling techniques, such as transformers or hierarchical models to further enhance the chatbot's ability to understand and respond effectively in complex and dynamic conversational contexts. Also, Develop mechanisms for continuous adaption and learning to evolving user preferences and language trends.To enable the chatbot to dynamically adjust its responses based on user feedback, this could involve incorporating reinforcement learning techniques .

References

[1] B. A. Shawar and E. Atwell, "Chatbots: are they really useful?" *Journal for Language Technology and Computational Linguistics*, vol. 22, no. 1, pp. 29–49, 2007.

[2] H. Bansal and R. Khan, "A review paper on human computer interaction," *Int. J. Adv. Res. Comput. Sci. Softw. Eng*, vol. 8, no. 4, p. 53, 2018.

[3] L. C. Klopfenstein, S. Delpriori, S. Malatini, and A. Bogliolo, "The rise of bots: A survey of conversational interfaces, patterns, and paradigms," in *Proceedings of the 2017 conference on designing interactive systems*, 2017, pp. 555–565.

[4] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.

[5] G. Bilquise, S. Ibrahim, and K. Shaalan, "Emotionally intelligent chatbots: A systematic literature review," *Human Behavior and Emerging Technologies*, vol. 2022, pp. 1–23, 09 2022.

[6] R. G. Athreya, A.-C. Ngonga Ngomo, and R. Usbeck, "Enhancing community interactions with data-driven chatbots–the dbpedia chatbot," in *Companion Proceedings of the The Web Conference 2018*, ser. WWW '18. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 2018, p. 143–146. [Online]. Available: https://doi.org/10.1145/3184558.3186964

[7] T.-H. K. Huang, J. C. Chang, and J. P. Bigham, "Evorus: A crowd-powered conversational assistant built to automate itself over time," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–13. [Online]. Available: https://doi.org/10.1145/3173574.3173869

[8] T. Hu, A. Xu, Z. Liu, Q. You, Y. Guo, V. Sinha, J. Luo, and R. Akkiraju, "Touch your heart: A tone-aware chatbot for customer care on social media," 2018.

[9] M. Jain, R. Kota, P. Kumar, and S. Patel, "Convey: Exploring the use of a context view for chatbots," 04 2018, pp. 1–6.

[10] F. Visin, K. Kastner, K. Cho, M. Matteucci, A. Courville, and Y. Bengio, "Renet: A recurrent neural network based alternative to convolutional networks," 2015.

[11] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," *arXiv preprint arXiv:1506.00019*, 2015.