EXPLORING ARTIFICIAL NEURAL NETWORK APPROACHES TO CONVERSATION MODELLING: A COMPREHENSIVE REVIEW

Raktim Dey, MDS202132
Rishika Tibrewal, MDS202135
Shreyansh Rastogi, MDS202144
Moumi Roy, MDS202125

May 6, 2023

ABSTRACT

Artificial neural network (ANN) approaches have been used extensively in recent years to model human conversation. Conversation modeling is a challenging task that requires the creation of models that can understand and generate natural language in a context-dependent way. ANN-based approaches have proven to be effective in this task, as they can capture the complex relationships between words, phrases, and sentences, and learn to generate coherent and contextually relevant responses. In this paper, we provide an overview of the various ANN-based approaches that have been used for conversation modeling, including recurrent neural networks (RNNs), transformer models, and memory-augmented neural networks. We discuss the advantages and limitations of these approaches, and highlight recent advances in the field, such as the use of pre-trained language models and the integration of multimodal information. We also discuss some of the challenges that remain in the field, including the need for more diverse and representative training data, the evaluation of models in real-world settings, and the development of models that can handle long-term dependencies and maintain consistency across multiple turns of conversation. Overall, ANN approaches to conversation modeling have the potential to revolutionize the way we interact with machines and provide a more natural and engaging user experience.

1 INTRODUCTION

Conversational agents, such as chatbots and virtual assistants, have become ubiquitous in our daily lives. They are used for a wide range of tasks, from customer service to personal assistance, and are becoming increasingly sophisticated in their ability to understand and respond to natural language. One of the key challenges in building conversational agents is modeling the complex and dynamic nature of human conversation, which involves understanding not only the meaning of individual words and sentences, but also the context, intentions, and emotions of the participants.

Artificial neural networks (ANNs) have been widely adopted in recent years for natural language processing (NLP) tasks, including conversation modeling. Two popular ANNs approaches for conversation modeling are Sequential Matching Networks (SMN) and Neural Conversational Models (NCM).

SMN is an ANN architecture specifically designed for response selection in conversational systems. It uses a dual-encoder architecture to encode both the question and candidate responses into a shared space and then measures the similarity between the two representations using various similarity metrics. SMN has been shown to achieve state-of-the-art results on benchmark datasets for response selection tasks.

In contrast, NCMs are a more general class of conversational models that can generate responses to a wide range of inputs, including questions, statements, and commands. NCMs can be trained on various types of data, including dialogues, chat logs, and user feedback, and they can be designed to incorporate additional information such as user preferences or contextual knowledge.

Both SMN and NCMs are examples of how ANNs can be applied to the challenging task of conversation modeling. This paper will provide a comparative analysis of these two approaches and discuss their strengths and weaknesses in terms of architecture, training, and application. The aim is to provide insights into the current state of ANNs approaches to conversation modeling and highlight the ongoing research and development in this exciting field.

2 LITERATURE SURVEY

The use of ANNs for conversation modeling has been a topic of research for several years. Initially, researchers used RNNs, which are capable of processing sequences of input data, to model conversation context and generate responses. For example, in the work of Serban et al. (2016), a sequence-to-sequence RNN was used to generate responses in a dialogue system. However, RNNs are limited in their ability to handle long-term dependencies and can suffer from the vanishing gradient problem, which makes it difficult to learn from long sequences. To address these limitations, Transformer Models, which are based on a self-attention mechanism, were introduced for conversation modeling. Transformer Models are capable of processing input sequences in parallel, which allows them to model long-term dependencies more effectively. In the work of Zhang et al. (2019), a Transformer-based model was used to generate responses in a task-oriented dialogue system.

More recently, Memory-Augmented Neural Networks (MANNs) have been used for conversation modeling. MANNs are designed to augment the network's memory with external memory banks, which can be used to store and retrieve information relevant to the conversation context. For example, in the work of Madotto et al. (2020), a MANN was used to store context information and generate responses in a multi-turn dialogue system.

Sequential Matching Networks (SMNs) are a Neural Network based approach that have been proposed to address the limitations of previous models. SMNs use multiple attention mechanisms to selectively attend to different parts of the input and output sequences and jointly learn representations of the input and output sequences. SMNs consist of two main components: an attention-based alignment module and a sequential matching module. The attention-based alignment module is responsible for aligning the two sequences and identifying the relevant parts of each sequence that are most important for the matching task. The sequential matching module then uses this aligned representation to compute a similarity score between the two sequences. In the work of Wu et al. (2019), SMNs were used to generate responses in a retrieval-based chatbot, achieving state-of-the-art performance on several benchmark datasets. In summary, the literature survey highlights the evolution of ANN approaches for conversation modeling, from RNNs to Transformer Models, MANNs, and SMNs. These models have been applied to various types of conversational agents, such as dialogue systems and chatbots, and have shown promising results. The integration of pre-trained language models has also improved the performance of these models. However, there are still challenges to be addressed, such as the need for more diverse and representative training data, and the development of models that can handle complex conversational dynamics.

3 Models and Algorithms

Sequential Matching Networks (SMNs) and Neural Conversational Models (NCMs) are two types of neural network architectures that are widely used in Natural Language Processing (NLP) tasks, such as text classification, question-answering, and conversational systems.

3.1 Models

3.1.1 Neural Conversational Models (NCMs):

- **Sequence-to-Sequence (Seq2Seq) Models:** These are a class of neural network architectures that are designed to map a variable-length input sequence to a variable-length output sequence. In NCMs, these are used to map the user's input sequence to the agent's output sequence.
- **Encoder-Decoder Models:** These are a class of Seq2Seq models that use two separate neural networks, an encoder and a decoder. The encoder network is used to encode the input sequence into a fixed-length vector, which is then passed on to the decoder network to generate the output sequence.

3.1.2 Sequential Matching Networks (SMNs):

- **Recurrent Neural Networks (RNNs):** RNNs are a class of neural networks that are designed to process sequential data. These use a feedback loop to allow information to persist over time, which makes them ideal for processing sequential data such as text. Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRUs) are generally used as the core building blocks in SMNs.
- **Convolutional Neural Networks (CNNs):** CNNs are neural networks that are designed to process images and other grid-like structures. In SMNs, these are used to extract features from the input sequence, which are then passed on to the RNN layers.
- **Attention Mechanisms:** These are used in SMNs to focus on specific parts of the input sequence that are relevant to the task at hand. A set of weights are calculated for each input element, which are then used to compute a weighted sum of the input elements.

3.2 Algorithms

3.2.1 Neural Conversational Models (NCMs):

- **Backpropagation:** This is an algorithm used to train neural networks by adjusting the weights of the network based on the error between the predicted output and the actual output. It is used to optimize the parameters of the NCM.

- **Teacher Forcing:** This is a training technique used in NCMs to improve the accuracy of the generated responses. The ground truth response is used as input to the decoder at each time step during training, instead of the generated response from the previous time step. This technique encourages the decoder to generate accurate responses by providing it with the correct input at each time step.
- **Beam Search:** This is a decoding algorithm used in NCMs to generate the most likely response given the input sequence. It maintains a list of the top K most likely candidate responses at each time step and prunes the list to keep only the K best candidates based on a scoring function. It is used to improve the fluency and coherence of the generated responses.
- **Reinforcement Learning:** It is a training technique used in NCMs to optimize the response generation process. The agent receives a reward or penalty signal based on the quality of the generated response. The agent then adjusts its response generation strategy to maximize the reward signal. It is used to improve the diversity and creativity of the generated responses.
- **Variational Autoencoders (VAEs):** VAEs are a class of generative models used in NCMs to learn a latent representation of the input sequence. VAEs consist of an encoder network, which maps the input sequence to a latent vector, and a decoder network, which maps the latent vector to the output response. VAEs are used to improve the coherence and relevance of the generated responses by learning a meaningful representation of the input sequence.

3.2.2 Sequential Matching Networks (SMNs):

- **Backpropagation:** It is an algorithm used to train neural networks by adjusting the weights of the network based on the error between the predicted output and the actual output. It is used to optimize the parameters of the SMN.
- **Stochastic Gradient Descent (SGD):** SGD is an optimization algorithm that is used to minimize the loss function of the SMN. In SGD, the parameters of the SMN are updated in the direction of the steepest gradient of the loss function.
- **Adam:** Adam is a popular optimization algorithm that combines the advantages of both SGD and RMSProp. It adapts the learning rate for each parameter based on the first and second moments of the gradients.

4 Results

Sequential matching networks (SMN) have been applied to several real-world datasets and have shown impressive performance in various text matching tasks. One example is the TREC QA dataset, which consists of questions and answers from newswire articles. SMN has achieved state-of-the-art performance on this dataset, with accuracy scores in the range of 70-80%, which is significantly higher than the performance of traditional information retrieval methods.
Another example is the WikiQA dataset, which consists of questions and answers from community-generated content on Wikipedia. SMN has also achieved state-of-the-art performance on this dataset, in terms of accuracy. SMN has also been applied to other text matching tasks, such as sentence matching and paraphrase identification. For example, on the SNLI dataset, which consists of pairs of sentences and requires determining whether the second sentence is entailed, contradicted, or neutral with respect to the first sentence, SMN has achieved accuracy scores of around 85%. Overall, SMN has demonstrated impressive performance on real-world datasets and has shown the ability to model the interaction between two sequences of text and capture subtle semantic relationships between them.

Neural conversational models (NCM) have been applied to various real-world datasets and have shown impressive performance in dialogue generation tasks. One example is the Persona-Chat dataset, which consists of conversations between two people, where each person has a predefined persona that the model must take into account when generating responses. NCMs have achieved near-human performance on this dataset, in terms of metrics such as perplexity and BLEU score. Another example is the Cornell Movie Dialogs Corpus, which consists of conversations between characters in movie scripts. NCMs have been applied to this dataset and have shown the ability to generate coherent and engaging responses that are consistent with the context of the conversation. NCMs have also been applied to other dialogue-generation tasks, such as customer service chatbots and personal assistants. For example, Google's Meena model, which is a type of NCM, has been shown to generate more human-like and engaging responses than previous dialogue models. Overall, NCMs have demonstrated impressive performance on real-world datasets and have shown the ability to generate coherent and engaging responses in a variety of dialogue-generation tasks.

5 Conclusion

Sequential matching networks (SMN) and neural conversational models (NCM) are two examples of ANN approaches that have shown impressive performance in different aspects of conversation modeling. SMN has been particularly successful in text matching tasks, such as information retrieval and question answering, while NCMs have excelled in dialogue generation tasks, such as chatbots and personal assistants.

Both SMN and NCMs have been applied to a variety of real-world datasets and have demonstrated their ability to generate coherent and engaging responses in different conversational contexts. Moreover, ongoing research in ANN approaches to conversation modeling promises to continue pushing the boundaries of what is possible in terms of natural and human-like interaction between humans and machines.

In conclusion, ANN approaches to conversation modeling, including SMN and NCMs, have the potential to significantly improve the way we interact with technology, making dialogue systems more human-like and engaging.

6 Usability of ChatGPT in writing this term paper

As a language model, ChatGPT is a powerful tool that can aid in the writing process for a term paper. With its ability to generate human-like text, provide suggestions for sentence structure, and offer a wealth of information on various topics, ChatGPT can significantly enhance the writing experience and improve the overall quality of the term paper.

ChatGPT can produce paragraphs of coherent and insightful content that can be used to supplement research or provide a starting point for further development. Moreover, it is able to generate grammatically correct sentences. It can offer suggestions for sentence structure and provide examples of how to rephrase awkward or unclear sentences, making the writing more polished and easier to understand. This can save the writer time by providing real-time feedback and reducing the need for multiple rounds of editing.

Additionally, it can be a valuable resource for researching and organizing information for a term paper. Providing access to vast amounts of information on various topics, it can help writers quickly find relevant sources and identify key points to include in the paper. It can also summarize complex information into more digestible formats, such as bullet points or outlines, making it easier to organize and present the information in the paper.

Despite its many benefits, there are some limitations to using ChatGPT for writing a term paper. While it can generate text and provide suggestions, ChatGPT is not a replacement for critical thinking or independent research. It should be used as a tool to supplement and enhance the writing process, rather than as a crutch to rely on for all aspects of the paper. Certain sections of the paper like models and algorithms, had to be re-generated multiple times to get the true and consistent information. While generating long sentences, ChatGPT got stuck in between and stopped abruptly, giving incomplete responses. When asked for research papers on a topic, the references and links generated did not exist. When "open source" papers were specifically mentioned, the results were appropriate.

In conclusion, the usability of ChatGPT for writing a term paper is significant. Its ability to generate text, provide suggestions for sentence structure, and offer a wealth of information on various topics can significantly enhance the writing experience and improve the overall quality of the paper. However, it is important to remember that ChatGPT should be used as a tool to supplement the writing process and not as a replacement for critical thinking or independent research.

7 References

- Sequential Matching Network: A New Architecture for Multiturn Response Selection in Retrieval-Based Chatbots by Yu Wu, Wei Wu, Chen Xing, Zhoujun Li, Ming Zhou.
- A Neural Conversational Model by Oriol Vinyals, Quoc Le.
- A Deep Relevance Matching Model for Ad-hoc Retrieval by Jiafeng Guo, Yixing Fan, Qingyao Ai, et al