

# Comparative study on personalized open-domain generative dialogue systems

**Yang Gu**

College of Artificial Intelligence, Beijing University of Posts and  
Telecommunications, Beijing, China

[gouy2@bupt.edu.cn](mailto:gouy2@bupt.edu.cn)

**Abstract.** Developing personalized dialogue systems is a monumental challenge in the rapidly evolving domain of Natural Language Processing (NLP). These systems are engineered to facilitate more natural, intuitive, and user-friendly interactions between humans and machines. Their theoretical significance extends to real-world applications, making them an invaluable asset in today's digital age. Over recent years, this area has garnered immense attention from researchers and academics, becoming a hotbed for innovation and study. This paper aims to contribute to this burgeoning field by conducting an in-depth comparative analysis of the current leading personalized language models. Specifically, this paper will scrutinize models grounded in deep learning algorithms and those that employ reinforcement learning techniques. This research objective is to dissect each approach's unique advantages and limitations. By doing so, the paper hopes to identify actionable avenues for improvement and optimization. Furthermore, the paper will offer a forward-looking perspective, outlining potential advancements and innovations that could shape the future landscape of personalized dialogue systems in Natural Language Processing.

**Keywords:** Deep Learning, Reinforcement Learning, Open-domain Dialogue Systems, Pre-trained Language Models.

## 1. Introduction

Internet products have become deeply integrated into people's daily lives in recent years. Accompanied by algorithmic advancements and increased computational power, deep learning technology has made significant progress in various aspects, including computational efficiency, parameter quantity, and expressive capabilities. Many researchers are applying deep learning technology to dialogue research and everyday life, with some products already in mass production, such as Microsoft's Xiaoice, Alibaba's intelligent customer service Xiaomi, and Apple's Siri. Particularly, the rise of large language models (LLMs) has brought ChatGPT into the public eye. People are amazed at the rapid progress of intelligent dialogue systems and can easily foresee their broad application prospects and research value.

Therefore, this paper will explore popular personalized language models through comparative methods, analyze their strengths and weaknesses, and propose feasible improvements to offer new perspectives and positive directions for exploration.

## 2. Project approach

In the rapidly evolving landscape of technology, dialogue systems have become an integral part of various industrial applications. These systems can be broadly classified into two distinct categories: task-oriented dialogue systems and open-domain dialogue systems, also known as casual conversation systems.

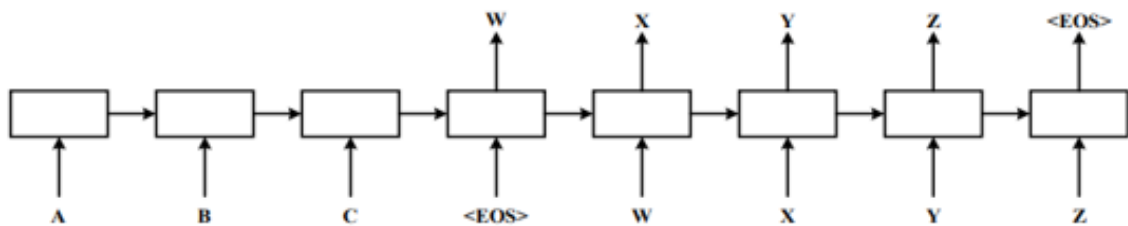
Task-oriented dialogue systems are designed with a specific purpose, usually to assist users in accomplishing particular tasks within a specialized domain. For example, these systems can help users book flights, make reservations, or navigate through customer service inquiries. They operate based on the user's input and aim to provide accurate and efficient services tailored to the task. A prime example of this type of system is Tmall Genie, an intelligent customer service platform that assists users in various specialized tasks.

Open-domain dialogue systems are more versatile and flexible on the other side of the spectrum. Unlike their task-oriented counterparts, these systems do not have a predefined purpose or topic limitations. This allows for a more free-flowing and natural interaction between the machine and the user. Voice assistants like Xiao Ai and advanced language models like ChatGPT fall under this category. These systems offer a more personalized and engaging user experience, as they can adapt to various topics and conversational styles.

Given the increasing demand for more personalized and dynamic interactions, this report strongly emphasizes the research and exploration of open-domain dialogue systems. The aim is to delve deeper into these systems' capabilities, limitations, and future potential to enhance user engagement and satisfaction.

There are mainly two methods for building open-domain dialogue systems: one is based on retrieval, and the other is based on generation. The retrieval-based method first constructs a search corpus and then uses a semantic matching model to calculate the similarity score between the user's input and the responses. Based on this score, an appropriate response is selected as the final result. Since these responses are extracted from the corpus, they are generally fluent. However, these responses are highly dependent on the existing sentences in the corpus and cannot generate new replies. Therefore, this method has significant limitations in open-domain dialogue systems.

On the other hand, the generative approach involves incorporating training data from multiple topics during the training process. The dialogue sequences are then fed into a seq2seq model to predict responses (as shown in Figure 1) [1]. This method can generate sentences that have never appeared in the corpus, making it more suitable for open-domain human-machine dialogue modeling. However, its drawback is that the model often generates meaningless answers due to training samples like "I don't know" and "Okay."



**Figure 1.** Classic seq2seq framework structure [1].

Many researchers have proposed different methods to address this issue from various angles. For example, some use entropy-based algorithms to filter out generic responses from the dataset. In contrast, others use memory modules to extract and store helpful information from the training corpus, among other methods, all aimed at reducing meaningless replies.

This study focuses on two types of methods: 1) Based on deep learning, it uses end-to-end neural dialogue models to learn character role information from large-scale datasets, thereby giving open-domain robots personality and improving the diversity and personalization level of the dialogue. 2)

Based on reinforcement learning, it continuously optimizes the quality of generated replies during user interactions and adjusts reply strategies based on user feedback, thus generating more natural and appropriate replies. By comparing the experimental results, the study aims to determine which direction is more effective for open-domain dialogue systems. The following is a detailed experimental introduction.

### 3. Experimental analysis

#### 3.1. Deep learning

For the first type of method mentioned above, this study focuses on an improved personalized language model based on the Transformer, called IMDPchat [2]. It mainly consists of four parts:

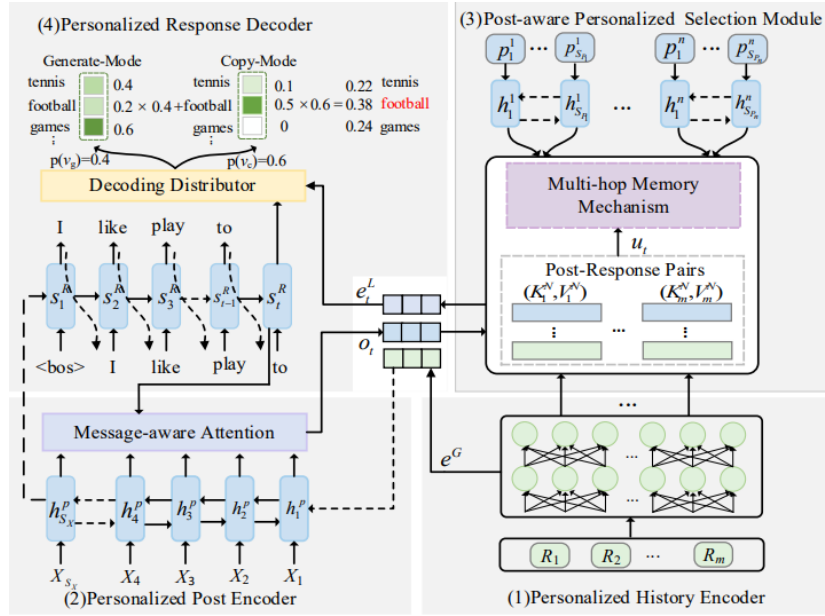
(1) Personalized History Encoder and Global User Role Information: Utilizes a personalized language model based on the Transformer to encode historical replies and constructs global user role information based on the representation of dialogue history replies.

(2) Personalized Post Encoder: Uses BiGRU to encode the input context into a personalized context representation [3].

(3) Post-Aware Personalized Selection Module: Adopts a custom multi-hop memory mechanism to calculate the relevance between the current and historical dialogue contexts. Then, highly relevant historical dialogue replies are selected from the key-value memory network, and dynamic user role information is constructed based on the corresponding historical dialogue replies.

(4) Personalized Response Decoder: Predicts replies during the decoding phase based on the generated personalized representation, global user role information, and dynamic user role information.

Figure 2 is a schematic diagram of the various components of this model.



**Figure 2.** IMDP model structure diagram [2].

Regarding the training aspect of the dialogue model, the primary objective is to optimize the model so that it can most accurately generate the desired or target response based on the user's immediate context and historical dialogue data. In technical terms, the model aims to maximize the likelihood of producing the correct output reply given the input variables, including the current conversational context and past interactions with the user.

A specific loss function is employed to quantify the model's performance during the training phase. This loss function denoted as:  $L = -\sum_{t=1}^{S_y} \log[p(y_t|y_{<t}, X, H)] - \phi S_y$ , serves as a mathematical

representation of the difference between the model's predictions and the actual target replies. The goal is to minimize this loss function, making the model's generated responses increasingly aligned with the expected or target replies. Doing so makes the model more effective and reliable in generating contextually appropriate and accurate responses in real-time conversations.  $\Phi$  : A specific hyperparameter is employed to regulate the penalty associated with the length of the generated output. This allows for fine-tuning the model's responses to be either more concise or detailed, depending on the context.

$p(y_t|y_{<t}, X, H)$ : The likelihood of producing a particular word, denoted as  $y_t$ , is determined by considering both the input context and the historical interactions with the user. This probability guides the model in generating the most appropriate word in a given conversational scenario.

### 3.2. Reinforcement Learning

For the second type of method mentioned above, reinforcement learning is a machine learning approach whose main goal is to enable intelligent agents to learn how to maximize reward signals through interaction with the environment. In dialogue systems, reinforcement learning can be used to learn the optimal dialogue strategy to improve the quality and efficiency of the conversation.

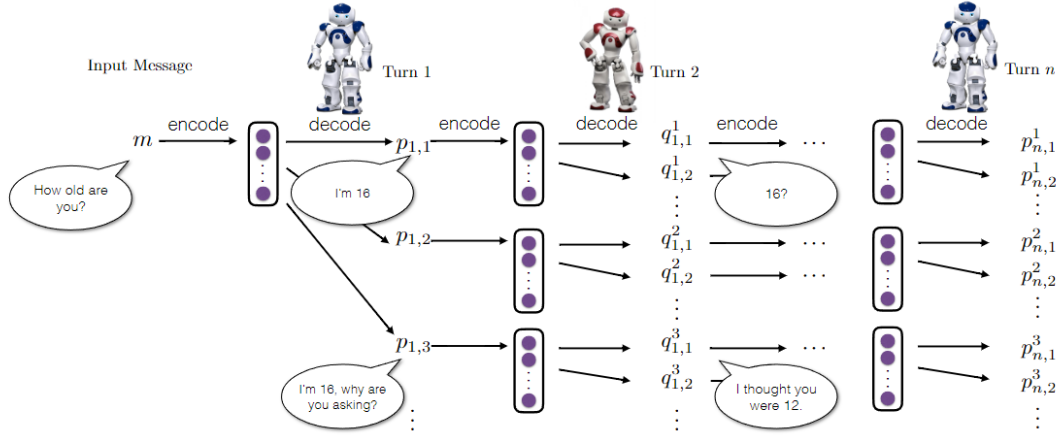
The design of dialogue strategies based on reinforcement learning usually involves three main components: state representation, action selection, and reward design. State representation is the process of converting the dialogue state into a form that the computer can process for handling within the intelligent agent. Action selection is the process by which the intelligent agent chooses the optimal action in a given state. Reward design is the process of defining the reward signal so that the intelligent agent can learn the optimal dialogue strategy. Figure 3 shows a dialogue simulation between two agents.

(1) Action ( $A_t$ ): The action, denoted as  $A_t$ , involves generating a response in the dialogue, which is formulated as  $A_t = \text{gen}(S)$ . Here,  $\text{gen}(S)$  refers to the sequence outputted by the Decoder. The scope of possible actions is limitless, as the Agent can produce sentences of varying lengths.

(2) State ( $S_t$ ): By inputting the information that must generate a response into the current dialogue box (state), the dialogue box is converted into a vector representation.

(3) Policy: The policy is represented as  $P_{RL}(P_{i+1}|P_i)$ , where  $P_{i+1}$  is the reply generated in the context of a specific dialogue. In this setup, the policy is the probabilistic mapping between a given dialogue state and the subsequent actions, which are also dialogues. This means that the policy dictates the likelihood of generating specific responses based on the current state of the conversation.

(4) Reward: The model incorporates three distinct internal rewards to address the challenges of generating natural language in seq2seq frameworks. These are Responsiveness ( $r_{EA}$ ), which gauges how adequately the generated text answers the user's question; Logical Consistency ( $r_{SC}$ ), which assesses the rationality and relevance of the generated dialogue; and Emotional Awareness ( $r_{EI}$ ), which evaluates the emotional tone and appropriateness of the response. These internal incentives guide the model in creating more meaningful, contextually relevant, and emotionally sensitive replies.



**Figure 3.** Dialogue simulation between the two agents [4].

### 3.3. Dataset

The research employs data from two online discussion platforms: Weibo and Reddit, some indicators of the dataset are shown in Table 1. The Weibo data is a specialized subset of a larger dataset known as PChatbotW, encompassing a year's worth of Weibo posts starting September 10, 2018 [5].

Weibo is a widely-used social networking site in China, where users can share brief messages, commonly known as "contexts." These messages can receive replies from other users, and each message and reply is tagged with a unique user ID for identification. On the other hand, the Reddit data is extracted from posts made on the Reddit online forums from December 1, 2015, to October 30, 2018. Unlike Weibo, Reddit conversations can form hierarchical, tree-like structures of replies. This allows for exploring these tree structures to pair each parent comment with its subsequent child comments, thereby creating various sets of historical conversational contexts and their corresponding replies.

**Table 1.** Some indicators of the dataset.

	Weibo	Reddit
Number	420000	280642
Avg.history length	32.3	85.4
Avg.length of post	24,9	10.5
Avg.length of response	10.1	12.4
Number of training samples	3000000	2000000
Number of validation samples	600000	403210
Number of testing samples	600000	403210

### 3.4. Comparative models

To evaluate these two models' effectiveness, this study compared several representative baseline models, which can be categorized into four types.

(1) Non-personalized generative models:

Seq2SeqWA: This is a Seq2Seq model based on GRU and equipped with an attention mechanism [6].

MMI: This is also a Seq2SeqWA model. However, it uses Maximum Mutual Information as the new objective function to reduce the proportion of generic responses, thus generating more diverse and interesting replies [7].

(2) Personalized generative models using user role vectors:

Speaker: This model based on Seq2SeqWA inputs user role vectors into the decoder to assist in dialogue generation [8].

PersonaWAE: This personalized Wasserstein autoencoder uses learned user role vectors as conditions to construct a Gaussian mixture distribution [9].

(3). Personalized generative models using predefined role information:

PerWAE: This model employs a memory-enhanced architecture to mine character roles in the context and combines it with a conditional variational autoencoder model to generate more diverse responses [10].

PEE: This model mines and associates existing dialogue corpora using a topic model based on variational autoencoders [11].

(4). Personalized generative models using implicit role information:

VHRED-P: This is a latent variable hierarchical recurrent encoder-decoder model trained by maximizing the variational lower bound on the log-likelihood [12].

ReCoSa-P: This model uses self-attention mechanisms to calculate attention weights between each dialogue context and response representation [13].

### 3.5. Evaluation metrics

(1) BLEU-1, BLEU-2, ROUGE-L: These are evaluation metrics used to measure the word-level accuracy between the generated responses and the ground truth labels. Higher values indicate higher similarity between the generated and actual responses [14, 15].

(2) Dist-1/2: This is an important metric for evaluating the performance of text generation models. It assesses the similarity between the generated text and the reference text. A higher Dist-1/2 value means the generated responses have greater diversity [16].

(3) Greedy Matching: This uses a metric method based on embeddings to calculate the semantic relevance between the generated and actual responses through cosine similarity. This helps assess the responses' quality and similarity [17].

## 4. Results and discussion

### 4.1. Experimental results

The experimental results are shown in Table 2.

**Table 2.** Experimental results.

Dataset	Model	BLEU-1	BLEU-2	ROUGE-L	Dist-1	Dist-2	Greedy
Weibo	Scq2SeqWA	3.332	0.289	8.742	0.940	2.187	0.258
	MMI	3.635	0.095	5.315	9.714	42.479	0.303
	Speaker	4.987	0.203	7.983	5.132	19.112	0.311
	PersonaWAE	3.506	0.149	10.305	2.489	19.713	0.307
	PerCVAE	5.115	0.299	7.956	14.075	49.741	0.295
	PEE	6.526	0.692	8.123	10.923	31.256	0.318
	VHRED-P	6.989	0.712	10.789	2.201	7.847	0.309
	ReCoSa-P	7.315	0.829	12.562	1.652	4.458	0.311
	IMDPchat	<b>9.254</b>	<b>0.896</b>	<b>13.698</b>	15.272	55.704	0.315
Reddit	Reinforcement	8.734	0.819	12.395	<b>18.482</b>	<b>60.873</b>	<b>0.324</b>
	Scq2SeqWA	1.819	0.023	4.068	5.203	19.485	0.472
	MMI	2.065	0.011	3.792	5.914	31.093	0.454
	Speaker	2.642	0.05	4.523	8.951	34.187	0.457
	PersonaWAE	2.637	0.112	7.856	1.758	25.917	0.442
	PerCVAE	5.879	0.576	8.212	9.631	40.213	0.499
	PEE	5.825	0.612	8.325	6.217	29.256	0.512
	VHRED-P	5.847	0.618	8.354	2.750	30.756	0.472
	ReCoSa-P	6.113	0.686	8.789	2.593	25.767	0.510
	IMDPchat	<b>6.898</b>	<b>0.689</b>	<b>11.852</b>	15.629	<b>64.927</b>	0.538
	Reinforcement	6.782	0.681	10.509	<b>17.928</b>	62.854	<b>0.570</b>

#### 4.2. *Experimental analysis*

From the above experimental data, it can be observed that in terms of evaluation criteria, both the improved personalized language model based on the Transformer architecture (IMDPchat) and the neural network models trained through reinforcement learning outperformed traditional baseline models. Furthermore, when comparing the test results of these two methods, it is clear that they perform similarly in terms of response accuracy, linguistic diversity, and semantic relevance, with little difference.

This indicates that both models have demonstrated excellent performance in generating personalized dialogues. This also means that significant progress has been made in this field, whether using end-to-end neural network methods or models trained through reinforcement learning. These achievements provide powerful tools for generating personalized dialogues and positively impact the field of natural language processing.

#### 5. Conclusion

Open-domain generative dialogue systems represent a widely applicable artificial intelligence technology that can simulate natural language conversations, offering enormous potential across various fields. However, traditional dialogue systems often require a large corpus for input to generate dialogue content, making them less adaptable to different scenarios and contexts. A more challenging issue is that they sometimes provide predictable, safe, and uncreative responses, making it difficult for users to experience refreshing conversations.

Moreover, the shortcomings of traditional dialogue systems also include the potential to overlook user input, lack of personalized responses, and limitations due to rigid generation methods. These issues indicate that existing personalized dialogue generation models still have various shortcomings in modeling dialogue generation tasks.

Therefore, this study aims to analyze their performance in generative dialogue systems by comparing two methods: an improved personalized language model based on the Transformer, and a neural network model trained through reinforcement learning. The evaluation metrics focused on in this study include the accuracy of responses, linguistic diversity, and semantic relevance. The research found that both methods can achieve better results in this field than previous models.

Furthermore, combining these two methods may yield even more outstanding results. A more powerful and innovative personalized dialogue generation system can be built by fully leveraging their strengths, providing users with a more attractive and satisfying conversational experience. This integrated approach is expected to lead the development of the generative dialogue system field, compensating for the shortcomings of traditional methods, and paving the way for further advancements in intelligent dialogue technology.

#### References

- [1] Li D, Chen Z, Cho E 2022 Overcoming catastrophic forgetting during domain adaptation of seq2seq language generation Proceedings of the 2022 Conference of the north American chapter of the association for computational linguistics: human language technologies 5441-5454
- [2] Yuanying W 2023 Research and application of Open Domain role-based dialogue Generation Method Shandong normal university
- [3] Chung J, Gulcehre C, Cho K H 2014 Empirical evaluation of gated recurrent neural networks on sequence modeling arXiv preprint arXiv: 1412.3555
- [4] Li J, Monroe W, Ritter A 2016 Deep reinforcement learning for dialogue generation arXiv preprint arXiv: 1606.01541
- [5] Chung J, Gulcehre C, Cho K H 2014 Empirical evaluation of gated recurrent neural networks on sequence modeling arXiv preprint arXiv: 1412.3555
- [6] Bahdanau D, Cho K, Bengio Y 2015 Neural machine translation by jointly learning to align and translate Proceedings of the international conference on learning representations

- [7] Chan Z, Li J, Yang X 2019 Modeling personalization in continuous space for response generation via augmented wasserstein autoencoders Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (emnlp-ijcnlp) 1931-1940
- [8] Zheng Y, Zhang R, Huang M 2020 A pre-training based personalized dialogue generation model with persona-sparse data Proceedings of the AAI conference on artificial intelligence **34** 05 9693-9700.
- [9] Onishi T, Shiina H 2021 Multi-turn dialogue generation considering speech information of each speaker by adding recurrent neural networks 2021 10th international congress on advanced applied informatics (IIAI-AAI) IEEE 524-527
- [10] Zhang H, Lan Y, Pang L 2019 Recosa: Detecting the relevant contexts with self-attention for multi-turn dialogue generation arXiv preprint arXiv: 1907.05339
- [11] Xu M, Li P, Yang H 2020 A neural topical expansion framework for unstructured persona-oriented dialogue generation arXiv preprint arXiv: 2002.02153
- [12] Serban I, Sordoni A, Lowe R 2017 A hierarchical latent variable encoder-decoder model for generating dialogues Proceedings of the AAI conference on artificial intelligence **31** 1
- [13] Papineni K, Roukos S, Ward T 2002 Bleu: a method for automatic evaluation of machine translation Proceedings of the 40th annual meeting of the Association for Computational Linguistics 311-318
- [14] Qian H, Li X, Zhong H 2021 Pchatbot: A large-scale dataset for personalized chatbot Proceedings of the international ACM SIGIR conference on research and development in information retrieval 2470-2477
- [15] Li J, Galley M, Brockett C 2016 A persona-based neural conversation model Proceedings of the Annual meeting of the association of computational linguistics,
- [16] Liu C W, Lowe R, Serban I V 2016 How not to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation arXiv preprint arXiv: 1603.08023
- [17] Zhang Y, Sun S, Galley M 2019 Dialogpt: Large-scale generative pre-training for conversational response generation arXiv preprint arXiv: 1911.00536