Improving Conversational Recommendation Systems via Counterfactual Data Simulation

Xiaolei Wang[†]
Gaoling School of Artificial Intelligence, Renmin University of China
Beijing, China
wxl1999@foxmail.com

Kun Zhou[†]
School of Information, Renmin University of China
Beijing, China
francis_kun_zhou@163.com

Xinyu Tang[†]
Gaoling School of Artificial Intelligence, Renmin University of China
Beijing, China
txy20010310@163.com

Wayne Xin Zhao[†✉]
Gaoling School of Artificial Intelligence, Renmin University of China
Beijing, China
batmanfly@gmail.com

Fan Pan
Huawei Poisson Lab
Beijing, China
panfan3@huawei.com

Zhao Cao
Huawei Poisson Lab
Beijing, China
caozhao1@huawei.com

Ji-Rong Wen[†]
Gaoling School of Artificial Intelligence, Renmin University of China
Beijing, China
jrwen@ruc.edu.cn

ABSTRACT

Conversational recommender systems (CRSs) aim to provide recommendation services via natural language conversations. Although a number of approaches have been proposed for developing capable CRSs, they typically rely on sufficient training data for training. Since it is difficult to annotate recommendation-oriented dialogue datasets, existing CRS approaches often suffer from the issue of *insufficient training* due to the scarcity of training data.

To address this issue, in this paper, we propose a **C**ounter**F**actual data simulation approach for **CRS**, named **CFCRS**, to alleviate the issue of data scarcity in CRSs. Our approach is developed based on the framework of *counterfactual data augmentation*, which gradually incorporates the rewriting to the user preference from a real dialogue without interfering with the entire conversation flow. To develop our approach, we characterize user preference and organize the conversation flow by the entities involved in the dialogue, and design a multi-stage recommendation dialogue simulator based on a conversation flow language model. Under the guidance of the learned user preference and dialogue schema, the flow language model can produce *reasonable, coherent* conversation flows, which can be further realized into complete dialogues. Based on the simulator, we perform the intervention at the representations of the interacted entities of target users, and design an adversarial training method with a curriculum schedule that can gradually optimize the data augmentation strategy. Extensive experiments show that our approach can consistently boost the performance of several competitive CRSs, and outperform other data augmentation methods, especially when the training data is limited. Our code is publicly available at https://github.com/RUCAIBox/CFCRS.

CCS CONCEPTS

• **Information systems** → **Recommender systems**.

KEYWORDS

Conversational Recommender System; Counterfactual Data Augmentation; Prompt Learning

ACM Reference Format:
Xiaolei Wang[†], Kun Zhou[†], Xinyu Tang[†], Wayne Xin Zhao[†✉], Fan Pan, Zhao Cao, and Ji-Rong Wen[†]. 2023. Improving Conversational Recommendation Systems via Counterfactual Data Simulation. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23), August 6–10, 2023, Long Beach, CA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3580305.3599387

1 INTRODUCTION

The recent success of conversational intelligence [1, 7] has empowered a more convenient way for information seeking by *conversational recommender systems* (CRSs) [3, 6, 14], which aims to provide high-quality recommendation service through multi-turn natural

† Beijing Key Laboratory of Big Data Management and Analysis Methods.
✉ Corresponding author.

KDD '23, August 6–10, 2023, Long Beach, CA, USA

ACM ISBN 979-8-4007-0103-0/23/08…$15.00
https://doi.org/10.1145/3580305.3599387

language conversations. Typically, a CRS recommends the suitable items that satisfy the user need via *a recommender module*, and generates the proper response based on the conversation context and predicted items via *a conversation module*. These two modules are systematically integrated to fulfill the information-seeking task.

To develop capable CRSs, various approaches have been proposed in the literature [19, 30, 50] based on deep neural networks. In particular, the powerful Transformer network [2] and pre-trained language models (PLM) [36] have largely raised the performance bar on conversational recommendation. These approaches rely on high-quality recommendation-oriented conversation data for model training. While it is difficult to manually create large-scale CRS datasets, which require well-trained annotators to generate *coherent, diverse* conversation flow in an information-seeking scenario [22, 54]. Therefore, existing CRS datasets [18, 22] are often limited in data size, lacking sufficient coverage of diverse information needs and user preferences. To alleviate this issue, existing studies incorporate external data (*e.g.,* knowledge graphs [2, 52]) and model resources (*e.g.,* DialoGPT [48]) to reduce the demand for training data in developing a capable CRS.

Despite the performance improvement, the fundamental issue of *insufficient training* in existing CRSs has not been well addressed due to the scarcity of training datasets. As a general solution to data scarcity, data augmentation techniques [5, 34, 43] have been widely applied in a variety of tasks, which either use heuristic strategies [40, 42] or learnable models [31, 45] for enlarging the data size. However, it is challenging to augment high-quality recommendation-oriented dialogues (short as *recommendation dialogues*) with automatic approaches, since it needs to mimic the interactive information-seeking process via a reasonable, coherent conversation flow. To be *reasonable*, the conversation scenario should be designed with meaningful user needs and suitable item recommendations, which conform to the factual information in the given domain. To be *coherent*, the augmented user preference should be consistent throughout the whole conversation, which should be well clarified and maintained as the conversation progresses. Considering these difficulties, existing work [13, 16] that uses specific rewriting strategies cannot generate high-quality CRS datasets.

To enhance the reasonableness and coherence of the conversation flow, we take a holistic perspective to develop the augmentation approach for recommendation dialogues by gradually incorporating the rewriting or adaptation into a real dialogue. Specially, each rewriting is expected to be carefully controlled without interfering with the entire conversation flow. Indeed, such intuition can be well fit into the framework of *counterfactual data augmentation* [8, 20, 27], which incorporates counterfactual learning for augmenting the limited data. In this setting, the essence of our approach is to answer the key question: "*What the dialogues would be if we intervene on the observed user preference?*", where user preference is considered to be the most important factor to determine a conversation flow. To instantiate it, we consider characterizing user preference and organizing the conversation flow by the entities involved in the dialogue (*e.g.,* movie actors and genres). Further, our rewriting strategy is implemented by a learnable edit function, which can produce informative edits to the entity representations for improving the recommendation module. In this way, the original user preference is gradually revised and finally reaches the level that a high-quality yet different conversation is augmented.

To this end, in this paper, we present the proposed **C**ounter**F**actual data simulation approach for **CRS**, named **CFCRS**, for alleviating the issue of data scarcity in CRSs. Our core idea is to leverage counterfactual learning to augment user preference and then employ the augmented user preference to simulate conversation data. Specifically, we design a recommendation dialogue simulator that can generate *reasonable, coherent* conversations for two target users. To guarantee the quality of the simulated conversations, we design a flow language model to generate the conversation flow, which is guided by the learned user preference and dialogue sketch. Based on the dialogue simulator, we perform the intervention at the representations of the interacted entities of target users, and design an adversarial training method with a curriculum schedule that can gradually optimize the edit function towards an improved recommendation capacity of CRSs.

To the best of our knowledge, it is the first time that *counterfactual data simulation* has been utilized to improve CRS models. Our proposed framework is agnostic to model implementations, hence is general to various CRS methods. To evaluate the effectiveness of our approach, we evaluate its performance with several representative CRS models on two public CRS datasets. Experimental results show that our approach can consistently boost the performance of these models, and outperform other data augmentation methods, especially when the training data is limited.

2 RELATED WORK

In this section, we summarize the related work as follows.

Conversational Recommender System. Conversational recommender systems (CRSs) aim to provide recommendation services through conversational interactions. One line of work [17, 33, 53] relies on pre-defined interactive actions (*e.g.,* asking preferences about item attributes or making recommendations) and hand-crafted templates to converse with users. Another line of work [2, 18, 52] focuses on interacting with users through more free-form natural language conversations. The above CRS methods are mostly developed by deep neural networks, which require sufficient high-quality data for training. However, it is expensive to annotate high-quality CRS examples and existing datasets are generally limited in scale. To address it, external resources like knowledge graphs [2, 52] and reviews [24, 55] have been introduced to enrich the datasets. However, the fundamental problem of insufficient training examples has not been well solved. In this work, we aim to solve the data scarcity problem via counterfactual data simulation.

Counterfactual Data Augmentation. Counterfactual data augmentation [20, 27, 56] focuses on generating unrecorded counterfactual examples from the real ones. Recently, it has been leveraged to alleviate the data scarcity problem and can improve the performance and robustness of deep neural networks. For example, in recommender systems, CASR [41] proposes to generate counterfactual user behavior sequences based on the real ones to supply the data for training sequential recommendation models. While for open-domain dialogue generation, CAPT [26] uses counterfactual

inference to automatically augment high-quality responses with different semantics to solve the one-to-many problem. In this work, we apply counterfactual data augmentation to the CRS task, aiming to obtain sufficient high-quality data. We also propose a curriculum learning strategy to gradually optimize the data augmentation strategy for CRSs.

3 APPROACH

In this section, we present the proposed **C**ounter**F**actual data simulation approach for **CRS**, named **CFCRS**, for alleviating the issue of data scarcity in CRSs, which is depicted in Figure 1.

3.1 Overview of Our Approach

Task Formulation. Conversational recommender systems (CRSs) aim to provide accurate item recommendation services through multi-turn natural language conversations. At each turn, the system either makes recommendations or chats with the user for preference elicitation. Such a process ends when the user accepts the recommended items or leaves. Formally, at the $(j+1)$-th turn, given the dialogue history $C_j = \{s_k\}_{k=1}^{j}$ consisting of j-turn utterances and the item set $\mathcal{I}$, the system should (1) select a set of candidate items $\mathcal{I}_j$ from the entire item set $\mathcal{I}$ to recommend, and (2) generate the response utterance s_{j+1} to the user. Besides, knowledge graph [2, 55] as an important auxiliary resource is usually available, denoted by $\mathcal{G}$. Typically, a CRS consists of the recommender module (parameterized by Θ_R) and the conversation module (parameterized by Θ_C), which are responsible for the recommendation and response generation tasks, respectively.

General Model Learning. Formally, let $\mathcal{D} = \{\langle C_j, r_j, i_j\rangle\}$ denote the set of training samples, where C_j is the dialogue history, r_j is the ground-truth response, and i_j is the recommended item for the j-th training sample. The optimization objectives for the two modules can be denoted as follows:

$$L_{\Theta_R}(\mathcal{D}) = -\sum_{\langle C_j, i_j\rangle \in D} \log g(i_j|C_j;\Theta_R), \tag{1}$$

$$L_{\Theta_C}(\mathcal{D}) = -\sum_{\langle C_j, r_j\rangle \in D} \log h(r_j|C_j;\Theta_C), \tag{2}$$

where $g(\cdot)$ and $h(\cdot)$ are the recommender and conversation modules, respectively. In the literature, existing CRSs mainly focus on designing various models or architectures to implement the two modules. For the recommendation module, it can be implemented with collaborative filtering [18], GNN [52], or Transformer [57] models. For the conversation module, it can be implemented with the vanilla Transformer [2] or PLM [39]. These approaches rely on high-quality CRS datasets to train the underlying models, which are often limited in size. To address this limitation, we propose to simulate high-quality data for conversation recommendation, which can be generally applied to various CRSs.

Counterfactual Learning for Dialogue Simulation. Our approach is inspired by the recent progress on *counterfactual data augmentation* [25], which incorporates counterfactual learning for augmenting the limited data. In our setting, the essence of our approach is to answer the core question: "*What the dialogues would be if we intervene on the observed user preference?*". As the basis of our approach, we design a recommendation-oriented dialogue simulator (short as *recommendation dialogue simulator*) that can generate *reasonable, coherent* conversations tailored for two target users (Section 3.2). Our recommendation dialogue simulator adopts a multi-stage generation process guided by the learned user preference and dialogue sketch: *flow schema → conversation flow → dialogue realization*. Based on the simulator, we construct the data augmentation via counterfactual learning (Section 3.3), which performs the intervention at the representations of the interacted entities of a target user. Further, we design an adversarial training method with a curriculum schedule that can gradually optimize the edit function towards an improved recommendation capacity of CRSs. In what follows, we introduce the two parts in detail.

3.2 Recommendation Dialogue Simulator

The goal of the recommendation dialogue simulator is to generate recommendation-oriented conversation data, so as to improve the performance of existing CRSs. Typically, it is difficult to create *fluent, coherent* conversation data, since it needs to simulate the free interaction for information seeking between two real users via chit-chat. As our solution, we develop a multi-stage generation process that first generates the *conversation flow* according to the predicted *flow schema* and then realizes the dialogue based on the generated flow. In our approach, we first introduce the basic concepts of conversation flow and flow schema.

3.2.1 Conversation Flow and Flow Schema. The conversation flow explicitly traces the key elements (*i.e.,* entities) of the information-seeking process. For example, given a two-turn conversation:

[Seeker]: *I love all kinds of comedy movies.*

[Recommender]: *Have you seen 21 Jump Street?*

[Seeker]: *Yes, I love this film because Jonah Hill is in it.*

[Recommender]: *Try another comedy movie with him, Superbad.*

we can derive a conversation flow: *comedy → 21 Jump Street → Jonah Hill → comedy → Superbad*. Based on such a flow, we can further generalize it into a flow schema: *genre → item → actor → genre → item*.

Formally, a conversation flow $f_{u,v}$ between two users u and v is characterized as a sequence of mentioned entities in a conversation arranged in the occurrence order, denoted by $f_{u,v} = \langle e_1, \cdots, e_j, \cdots, e_n\rangle$, and the corresponding flow schema (with an equal length to the flow) is characterized as a sequence of type tokens, denoted by $s_{u,v} = \langle t_1, \cdots, t_j, \cdots, t_n\rangle$, where e_i is a mentioned entity from a knowledge graph (KG) $\mathcal{E}$ and $t_i = \text{Type}(e_i)$ indicating the type of e_i. As we can see, conversation flow and flow schema are useful to generate concrete conversation content by capturing the entity preferences of users and establishing the dialogue sketch.

3.2.2 Preference Prompt Guided Flow Language Model. In our approach, we design a flow language model (FLM) that is parameterized by Θ_F based on the *preference prompts* for generating the conversation flow. Specifically, to generate a conversation flow, we first sample two target users u and v as the seeker and recommender, respectively, and then employ the two users to predict a flow schema $\hat{s}_{u,v}$. Then, the representations of target users (*i.e.,* user prompt) and the predicted schema (*i.e.,* schema prompt) are taken

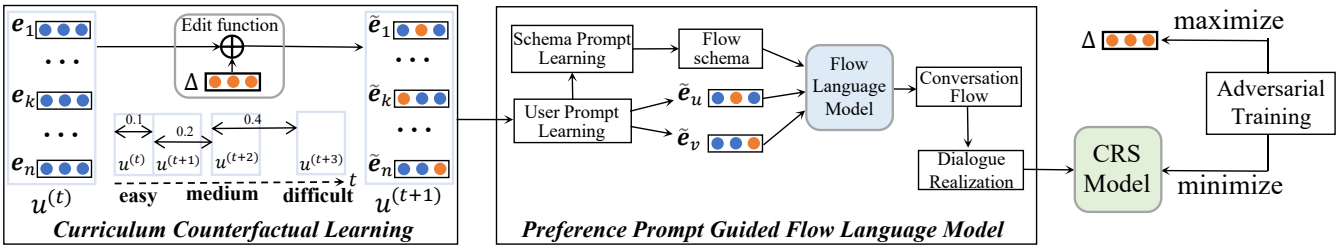

Figure 1: The overview of our approach CFCRS. We first adopt curriculum counterfactual learning to augment the user preference at the representation level, and then use the flow language model guided by user and schema prompts to generate conversation flows, which are then realized into dialogues. The edit function and CRS model are optimized with adversarial training to improve both the quality of the augmented data and the recommendation performance.

as the prompts to the FLM for generating the conversation flow as follows:

$$f_{u,v} \leftarrow \text{FLM}\left(\left[\underbrace{\boldsymbol{e}_u, \boldsymbol{e}_v}_{\text{user prompt}}, \underbrace{\{\boldsymbol{t}_j\}_{j=1}^{n}}_{\text{schema prompt}}\right]; \Theta_F\right). \tag{3}$$

Next, we discuss how to derive the two parts of prompts.

User Prompt Learning. Since the simulated dialogue occurs between the two target users, it is important to consider their preferences for generating the conversation flow. To capture the user preference, in recommender systems, it is common to assign each user a unique user ID, and learn the ID embedding based on the interacted items or entities as the user preference representation [12, 15]. However, in our simulation setting, we would like to generate more diverse user representations that are not limited to the real users in the CRS datasets. For this purpose, we do not explicitly maintain a user ID but learn ID-agnostic user representations. Specifically, following the previous work on CRSs [47, 52], we assume that a KG is available and extend this KG by attaching user nodes to their interacted entity nodes to compose a new heterogeneous knowledge graph (HKG), denoted as $\mathcal{G}$. To capture relational semantics between entities, we utilize R-GCN [32] to learn entity representations on the HKG. Formally, let n denote a node placeholder for the HKG, associated with an embedding vector $\boldsymbol{e}_n \in \mathbb{R}^{d_E}$ derived from R-GCN, where d_E denotes the embedding size. We utilize the self-attention mechanism to aggregate entity embeddings as the preference representation of the user u:

$$\begin{aligned} \boldsymbol{e}_u &= \mathbf{E}_u \cdot \boldsymbol{\alpha}, \\ \boldsymbol{\alpha} &= \text{softmax}(\boldsymbol{b}^\top \cdot \tanh(\mathbf{W}_\alpha \mathbf{E}_u)), \end{aligned} \tag{4}$$

where $\mathbf{E}_u$ is the matrix consisting of the embeddings of all the interacted entities of user u, $\boldsymbol{\alpha}$ is the attention weight vector reflecting the importance of each entity, and $\mathbf{W}_\alpha$ and $\boldsymbol{b}$ are trainable parameters. A major advantage of this representation method is that it can be easily adapted to new users by modeling the entity preference based on the associated interaction records.

Schema Prompt Learning. In order to produce *fluent, coherent* conversations, we employ frequent flow schema to structure the conversation. Although the conversation flows can be very diverse, the frequent flow schemas for a CRS corpus are usually limited. Thus, we consider employing real CRS datasets to construct flow schemas with frequent pattern mining algorithms [9], and obtain a set of frequent flow schemas, denoted as $\mathcal{S}$. Then, the schema prediction task is cast as a classification problem over the schema set $\mathcal{S}$ based on user preference. Formally, we compute the prediction probability for a flow schema as follows:

$$\Pr(s_{u,v}|u,v) = \text{softmax}\big(\text{MLP}([\boldsymbol{e}_u, \boldsymbol{e}_v])\big), \tag{5}$$

where u and v are the two target users involved in the dialogue, and $s_{u,v}$ is the flow schema. In practice, first, we select the most probable flow schema according to Eq. (5). Then, we obtain the corresponding type embeddings by decomposing the predicted schema into type tokens $\{\boldsymbol{t}_j\}_{j=1}^{n}$, where each type token embedding $\boldsymbol{t}_j$ is obtained by looking up the type embedding table.

Flow Language Model Pre-Training. To model the conversation flow, we construct an FLM based on Transformer, which utilizes the encoder-decoder architecture. Following Eq. (3), the encoder takes the learned user prompt (*i.e.,* $\boldsymbol{e}_u$ and $\boldsymbol{e}_v$) and schema prompt (*i.e.,* $\{\boldsymbol{t}_j\}_{j=1}^{n}$) as input, and the decoder generates the conversation flow in an autoregressive manner based on the prompt. Formally, let $\boldsymbol{e}_j$ be the embedding of a token e_j (an entity) in the conversation flow $f_{u,v}$, and the probability of the flow to be generated is formulated as:

$$\begin{aligned} \Pr(f_{u,v}) &= \sum_{j=1}^{n} \Pr(e_j|e_0, \ldots, e_{j-1}) \\ &= \sum_{j=1}^{n} \text{softmax}\big(\mathbf{W}[\boldsymbol{e}_j; \boldsymbol{z}_{u,v}] + \boldsymbol{b}\big) \end{aligned} \tag{6}$$

where $\boldsymbol{z}_{u,v}$ is the encoding of prompt, $\mathbf{W}$ and $\boldsymbol{b}$ are trainable parameters. To pre-train the FLM, we require large amounts of conversation flows from various user pairs. Since the original CRS dataset is usually limited in conversation size, we consider generating pseudo conversation flows for pre-training. The basic procedure consists of three major steps: (i) we first randomly sample one flow schema from the frequent flow schema set $\mathcal{S}$; (ii) then, we sample an entity sequence from the constructed HKG $\mathcal{G}$ as the conversation flow according to the schema; (iii) finally, we randomly divide the entities in the sequence into two groups, which correspond to the entity preference of two users involving the conversation. When

a schema cannot lead to a reachable entity path, we continue to sample another schema.

3.2.3 Dialogue Realization. With the pre-trained FLM and entity preference of users, we can generate the corresponding conversation flows at a large scale. Next, we realize the generated conversation flows into recommendation dialogues.

Here, we adopt a simple template-based approach for dialogue realization. Specifically, we first collect the templates from observed dialogues by delexicalization, *i.e.,* substituting the mentioned entities with placeholders, *e.g.,* "*<genre>*" for "comedy". Then, the entities in conversation flows can be sequentially filled into these templates as new recommendation dialogues. For instance, an utterance "*I am in a mood for something scary*" would be converted to the template "*I am in a mood for something <genre>*". Since the focus of this work is to enhance the recommendation ability of CRSs instead of the general chit-chat ability, we do not adopt pre-trained dialogue models (*e.g.,* DialogGPT [48]) to realize these utterances. Our proposed method is simple yet effective to ensure fluency in language and faithfulness in conversation flow.

After building the recommendation dialogue simulator, we can employ it to generate simulated data with user preference representations as input (as will be used in Section 3.3).

3.3 Curriculum Counterfactual Learning

Although the above recommendation dialogue simulator can effectively enlarge the dataset by simulation, it is still limited to the actual users in existing datasets. In this part, we introduce a curriculum counterfactual learning approach that can learn to generate diverse data by augmenting the user preference representations.

3.3.1 Counterfactual User Preference Augmentation. Recall that in Section 3.2.2, we utilize a self-attentive mechanism to learn the user preference representation based on the interacted entities with Eq. (4). Formally, the set of interacted entities of user u is denoted as $\mathcal{E}_u = \{e_1, \ldots, e_i, \ldots, e_n\}$, and their embeddings are also aggregated as a matrix $\mathbf{E}_u = [\boldsymbol{e}_1, \ldots, \boldsymbol{e}_i, \ldots, \boldsymbol{e}_n]$.

In existing work [21, 28], discrete [21] and continuous [28] item-level edits have been explored for user data augmentation. To combine the merits of both approaches, we consider an entity-level edit to augment new user preferences. Specifically, we revise one entity embedding at each time with a specially designed edit function $f(\cdot)$, in which the entity selection is *discrete* and the embedding revision is *continuous*. Such a way can generate diverse user preferences while gradually incorporating controllable revisions. To instantiate the edit function, a number of model choices can be considered, *e.g.,* neural networks. However, we empirically find that it is difficult to optimize such an edit function, due to the lack of supervision signals in real datasets. Thus, we consider a simple yet effective edit function that directly adds a small disturbance vector $\boldsymbol{\Delta}_i$:

$$f(\boldsymbol{e}_i) = \boldsymbol{e}_i + \boldsymbol{\Delta}_i, \tag{7}$$

We denote all the disturbance vectors as Θ_E.

According to [8], such a simplified edit function is easier to learn and interpret, since samples near the decision boundary are usually discriminative in revealing the underlying data patterns. For each user u, we can perform the edit k times, so as to produce k different augmentations, each editing one specific entity embedding in $\mathbf{E}_u$ to generate the augmented user preference $\tilde{\boldsymbol{e}}_u$ (originally $\boldsymbol{e}_u$).

3.3.2 Adversarial Learning with Curriculum Schedule. As our edit function can be learned in a differentiable manner, we propose to use adversarial learning to enhance the informativeness of the augmented user preference representations. Intuitively, a more informative training instance tends to cause *a larger loss* in the recommendation accuracy, as such a user preference has not been well captured by the *current model*. Taking an adversarial learning perspective, the counterfactual edit function (parameterized by Θ_E) aims to maximize the loss of the recommender module (parameterized by Θ_R), while the recommender module aims to minimize its loss on the simulated data. In addition, we perform adversarial learning with a curriculum schedule to stabilize the optimization.

Adversarial Training. Let $\pi_{\Theta_E}(C|\tilde{\boldsymbol{e}}_u, \tilde{\boldsymbol{e}}_v)$ denote the recommendation dialogue simulator, which returns the probability of generating a dialogue C given the edited embeddings of two users $\tilde{\boldsymbol{e}}_u$ and $\tilde{\boldsymbol{e}}_v$ by Θ_E. The learning objective can be formulated as follows:

$$J^{\Theta_R^*, \Theta_E^*} = \min_{\Theta_R} \max_{\Theta_E} \mathbb{E}_{C \sim \pi_{\Theta_E}(\cdot|\tilde{\boldsymbol{e}}_u, \tilde{\boldsymbol{e}}_v)} [L_{\Theta_R}(C) - \lambda \cdot \parallel \Theta_E \parallel_2^2], \tag{8}$$

where $\Theta_E = \{\boldsymbol{\Delta}_u, \boldsymbol{\Delta}_v\}$ is the edit vectors for users u and v (Eq. (7)), $L_{\Theta_R}(C)$ is the loss of the recommendation module for the generated data C, and λ is the regularization weight for Θ_E. Note that Eq. (8) presents the optimization objective for a pair of users, which can be easily extended to all user pairs. To optimize the above objective (with two groups of parameters Θ_R and Θ_E), we can alternatively optimize each group of parameters by keeping the other group fixed. It is relatively straightforward to train the parameters of the recommender module (*i.e.,* Θ_R) by a standard recommendation loss (*e.g.,* cross-entropy loss [30]). However, it is infeasible to directly optimize the edit vectors in an end-to-end way, since it involves the generation of discrete conversation data. To tackle this issue, we adopt the classic REINFORCE algorithm [44] to update the parameters as follows:

$$\Theta_E = \Theta_E + \alpha \left(\sum_{t=1}^{T} L_{\Theta_R}(C_t) \nabla_{\Theta_E} \log \left(\pi_{\Theta_E}(C_t|\tilde{\boldsymbol{e}}_u, \tilde{\boldsymbol{e}}_v) \right) - 2\lambda \cdot \Theta_E \right), \tag{9}$$

where α is the learning rate and T conversations are sampled.

Curriculum Arrangement. In order to keep the training stable, we consider a curriculum learning approach that gradually increases the augmentation level: small variations are encouraged at the beginning of training, while larger variations can be gradually applied to enhance the model capacity. As shown in Eq. (8), we incorporate a controlling weight λ on Θ_E. To simulate counterfactual data in an *easy-to-difficult* process, we dynamically tune the augmentation level in each iteration. Specifically, we apply an annealing mechanism to regularize Θ_E with a shrinking λ:

$$\lambda = \rho \times \delta^{(k)}, \tag{10}$$

where ρ is the initial weight, δ is the decay ratio, and k is the current course. In this way, as the course gets more difficult, *i.e.,* the augmentation level increases, the CRS model can continually learn from diverse and informative training samples to improve its performance.

Algorithm 1: The training algorithm of our framework.

Input: The conversational recommendation dataset $\mathcal{D}$, HKG $\mathcal{G}$
Output: Parameters of the recommender module Θ_R and conversation module Θ_C in CRSs.

Pre-train the parameters of the recommendation dialogue simulator Θ_S with the union of real and pseudo data sampled from $\mathcal{G}$.
Pre-train the parameters of the recommender module Θ_R using the real dataset $\mathcal{D}$.
for $k = 1 \rightarrow N$ **do**
  Set the regularization parameter λ according to curriculum arrangement using Eq. (10).
  Optimize the parameters of edit function Θ_E by maximizing the loss of the recommender module using Eq. (9) and derive new user preference $\tilde{\boldsymbol{e}}$.
  Use new user preference $\tilde{\boldsymbol{e}}$ to simulate data C with the recommendation dialogue simulator by Eq. (3) and Eq. (4).
  Optimize the parameters of the recommender module Θ_R by minimizing its loss on simulated data C.
end
Optimize the conversation module Θ_C with the augmented data.
return Θ_R and Θ_C.

3.4 Parameter Learning

The parameters of our framework consist of four groups, namely the recommendation dialogue simulator Θ_S, the counterfactual edit function Θ_E, and the recommender module Θ_R and conversation modules Θ_C of the target CRS model. Algorithm 1 presents the training algorithm of our framework.

First of all, we pre-train the parameters of the recommendation dialogue simulator Θ_S with the union of real and pseudo conversation flow data using the cross-entropy loss. After pre-training, parameters Θ_S are fixed. Then, we perform curriculum counterfactual learning to augment new data for CRS learning. In each iteration, the parameters of the counterfactual edit function Θ_E and the recommender module of the target CRS model Θ_R are optimized via adversarial learning using Eq. (8). Specifically, we first learn the counterfactual edit function to maximize the loss of the recommender module, and then optimize the recommender module to minimize its loss on the simulated data. After the curriculum learning schedule, we optimize the parameters of the conversation module Θ_C with the union of simulated and real data.

4 EXPERIMENT

In this section, we first set up the experiments, then report the results and give a detailed analysis.

4.1 Experimental Setup

Datasets. To verify the effectiveness of our approach, we conduct experiments on two widely used English CRS datasets, *i.e.,* ReDial [18] and INSPIRED [10]. The ReDial dataset is an English CRS dataset about movie recommendations, and is constructed through crowdsourcing workers on Amazon Mechanical Turk (AMT). Similar to ReDial, the INSPIRED dataset is also an English CRS dataset about movie recommendations, but with a much smaller size. The statistics of both datasets are summarized in Table 1.

Table 1: Statistics of the datasets after preprocessing.

Dataset	#Dialogues	#Utterances	#Items
INSPIRED	1,001	35,811	1,783
ReDial	10,006	182,150	51,699

Baselines. Here we consider two major tasks for CRS evaluation, namely recommendation and conversation. For comparison, we select several representative methods (including both CRS models and adapted PLMs) tailored to each task.

• *BERT* [4]: It is a bidirectional PLM pre-trained on large-scale general corpora. We utilize the [*CLS*] token for recommendation.

• *GPT-2* [29]: It is an autoregressive PLM pre-trained on large-scale general corpora. We concatenate the historical utterances as inputs and take the generated text for response while using the representation of the last token for recommendation.

• *DialoGPT* [48]: It continues to pre-train GPT-2 on large-scale dialogue corpora. We use it in the same way as GPT-2.

• *ReDial* [18]: It is proposed along with the ReDial dataset, which includes a conversation module based on HRED [35] and a recommendation module based on a denoising auto-encoder [11].

• *KBRD* [2]: It introduces DBpedia to enhance the semantics of entities mentioned in the dialogues.

• *BARCOR* [37]: It proposes a unified framework based on BART, which tackles two tasks with a single model.

• *UniCRS* [39]: It designs knowledge-enhanced prompts based on DialoGPT to fulfill both tasks in a unified approach.

Among these baselines, BERT, GPT-2 and DialoGPT are PLMs, where BERT and GPT-2 are pre-trained on general corpora while DialoGPT is pre-trained on dialogue corpora. ReDial, KBRD, BARCOR, and UniCRS are CRS methods, where ReDial and KBRD use mentioned entities for recommendation, BARCOR utilizes dialogue texts for recommendation, and UniCRS makes use of both entities and texts for recommendation. To verify the generality of our framework, we apply it to KBRD, BARCOR, and UniCRS. To demonstrate the effectiveness of our framework, we compare it with several representative data augmentation methods.

• *EDA* [42]: It augments new examples by randomly performing edit operations, *i.e.,* replacement, insertion, swap, and deletion.

• *Mixup* [46]: It augments new examples in the continuous latent space by linear interpolations of input representations and labels of two random examples.

Evaluation Metrics. Following existing work [2, 18], we adopt different metrics to evaluate the recommendation and conversation tasks separately. For the recommendation task, following [2, 54], we use Recall@k, MRR@k, and NDCG@k (k=10,50). For the conversation task, following [52, 55], we adopt Distinct-n (n=2,3,4) to evaluate the diversity of the generated responses. Besides, following KGSF [52], we invite three annotators to score the generated responses from two aspects, namely *Fluency* and *Informativeness*. The range of scores is 0 to 2. For all the above metrics, we calculate and report the average scores on all test examples.

Implementation Details. We implement some baseline models based on the open-source toolkit CRSLab [51], which contains

Table 2: Results on the recommendation task. The best methods in each group are marked in bold. Numbers marked with * indicate that the improvement is statistically significant compared with the baseline (t-test with p-value < 0.05).

Datasets	ReDial						INSPIRED					
Models	Recall@10	Recall@50	MRR@10	MRR@50	NDCG@10	NDCG@50	Recall@10	Recall@50	MRR@10	MRR@50	NDCG@10	NDCG@50
ReDial	0.129	0.287	0.003	0.004	0.005	0.011	0.117	0.285	0.004	0.003	0.005	0.012
BERT	0.156	0.357	0.055	0.063	0.079	0.121	0.179	0.328	0.067	0.085	0.098	0.133
GPT-2	0.147	0.327	0.051	0.056	0.071	0.107	0.112	0.278	0.063	0.076	0.089	0.128
DialoGPT	0.173	0.361	0.062	0.068	0.089	0.135	0.125	0.247	0.059	0.081	0.092	0.120
KBRD	0.170	0.366	0.063	0.072	0.088	0.131	0.210	0.390	0.112	0.118	0.135	0.172
KBRD-EDA	0.174	0.371	0.068	0.077	0.093	0.136	0.180	0.364	0.088	0.094	0.109	0.146
KBRD-mixup	0.189	0.390	0.072	0.081	0.099	0.144	0.210	0.390	0.104	0.113	0.122	0.165
KBRD-CFCRS	**0.206***	**0.408***	**0.084***	**0.093***	**0.109***	**0.156***	**0.226***	**0.426***	**0.123***	**0.129***	**0.145***	**0.188***
BARCOR	0.169	0.374	0.063	0.073	0.088	0.133	0.185	0.339	0.080	0.087	0.104	0.137
BARCOR-EDA	0.179	0.395	0.067	0.077	0.093	0.140	0.210	0.390	0.102	0.109	0.127	0.166
BARCOR-mixup	0.169	0.363	0.061	0.070	0.086	0.129	0.139	0.344	0.070	0.081	0.086	0.132
BARCOR-CFCRS	**0.198***	**0.406***	**0.079***	**0.088***	**0.107***	**0.151***	**0.246***	**0.421***	**0.114***	**0.122***	**0.145***	**0.183***
UniCRS	0.217	0.428	0.088	0.096	0.118	0.163	0.272	0.441	0.156	0.164	0.184	0.224
UniCRS-EDA	0.167	0.357	0.068	0.077	0.091	0.133	0.295	0.451	0.132	0.165	0.186	0.220
UniCRS-mixup	0.206	0.394	0.073	0.088	0.116	0.158	0.246	0.426	0.153	0.164	0.182	0.219
UniCRS-CFCRS	**0.231***	**0.444***	**0.096**	**0.111***	**0.129***	**0.175***	**0.308***	**0.466***	**0.168***	**0.176***	**0.204***	**0.242***

comprehensive CRS models and benchmark datasets. For the recommendation dialogue simulator, we adopt a Transformer with 12-layer encoders and decoders for the FLM, and its hidden size and embedding size are 768. To be consistent with the FLM, the hidden size of the user and schema prompt is also 768. In curriculum counterfactual learning, the maximum training iterations are set to 20, and we adopt the early stop strategy. The initial value of the regularization weight and the decay ratio are tuned in the range of $[10^{-1}, 10^{-2}, 10^{-3}]$ and $[0.9, 0.8, 0.7]$ for different models. We use AdamW [23] with the default parameter setting to optimize the model parameters. Its learning rate is mostly set to $1e^{-4}$ and tuned in the range of $[5e^{-5}, 1e^{-4}, 5e^{-4}, 1e^{-3}]$.

4.2 Evaluation on Recommendation Task

In this part, we conduct experiments to evaluate the effectiveness of our model on the recommendation task.

Automatic Evaluation. Table 2 shows the performance of different methods on the recommendation task. First, we can see that KBRD, BARCOR, and UniCRS mostly outperform the other baselines in all metrics. The three methods all incorporate external KGs to enrich the information of mentioned entities in the conversation context, which can effectively alleviate the data scarcity problem and better capture user intents and preferences. Among the three methods, UniCRS performs the best in all metrics. UniCRS utilizes knowledge-enhanced prompts to guide the PLM, and incorporates a pre-training task to improve the quality of prompts. Such a way can effectively endow the PLM with entity knowledge for better performance.

Second, for the two data augmentation baselines, we observe that most of the time they both improve the performance of the three CRS methods. It indicates the effectiveness of data augmentation strategies in the CRS task, since the training data is not sufficient. However, we can see that the improvement is not stable, and even causes performance degradation for UniCRS on the ReDial dataset. A possible reason is that the two methods only rely on heuristic rules to modify the original examples for augmenting new ones, which makes it hard to guarantee the quality of the augmented data and may even produce abnormal conversations.

Table 3: Ablation analysis on the recommendation task. "FLM" denotes the flow language model and "Template" denotes template-based dialogue realization. "-" denotes removing the corresponding component.

Datasets	ReDial		INSPIRED	
Metrics	Recall@10	Recall@50	Recall@10	Recall@50
BARCOR	0.169	0.374	0.185	0.339
+CFCRS	**0.198**	**0.406**	**0.246**	**0.421**
-Curriculum	0.187	0.399	0.190	0.405
-Adversarial	0.184	0.389	0.225	0.395
-FLM	0.181	0.385	0.211	0.390
-Frequent Schema	0.186	0.394	0.246	0.407
-Template	0.183	0.382	0.174	0.385

Finally, we can see that our model can improve the performance of the three CRS methods by a large margin. It indicates the effectiveness and generality of our framework. Furthermore, our approach mostly outperforms the two data augmentation baselines significantly. In our approach, we use a counterfactual data simulation approach, which includes a pre-trained FLM to guarantee the coherence of the conversation flow and adversarial training to enhance the informativeness of simulated data. Besides, we utilize the curriculum learning strategy to gradually optimize CRS models using examples with different augmentation levels, which further improves the stability of the training process.

Ablation Study. Our approach incorporates several components to improve the quality of the augmented data. To verify the effectiveness of each component, we conduct the ablation study on BARCOR using the ReDial and INSPIRED datasets, and report the results of Recall@10 and Recall@50. The results are shown in Figure 3. We

Table 4: Automatic evaluation results on the conversation task. We abbreviate Distinct-2,3,4 as Dist-2,3,4. The best methods in each group are marked in bold. Numbers marked with * indicate that the improvement is statistically significant compared with the baseline (t-test with p-value < 0.05).

Datasets	ReDial			INSPIRED		
Models	Dist-2	Dist-3	Dist-4	Dist-2	Dist-3	Dist-4
ReDial	0.023	0.236	0.228	0.153	0.255	0.397
GPT-2	0.354	0.486	0.441	2.347	3.691	4.568
DialoGPT	0.476	0.559	0.486	2.408	3.720	4.560
KBRD	0.198	0.339	0.473	0.223	0.415	0.616
KBRD-EDA	0.323	0.476	0.565	0.466	0.856	1.174
KBRD-mixup	0.172	0.292	0.449	0.362	0.680	0.987
KBRD-CFCRS	**0.477***	**0.603***	**0.728***	**0.573***	**1.148***	**1.645***
BARCOR	0.404	0.540	0.654	2.923	4.172	4.992
BARCOR-EDA	0.522	0.698	0.717	3.597	5.108	5.959
BARCOR-mixup	0.568	0.704	0.740	3.856	5.582	6.576
BARCOR-CFCRS	**0.701***	**0.971***	**0.969***	**4.081***	**5.953***	**6.979***
UniCRS	0.351	0.631	0.897	2.809	4.530	5.555
UniCRS-EDA	0.440	0.801	1.141	2.882	4.859	6.166
UniCRS-mixup	0.412	0.701	0.918	2.517	4.173	5.496
UniCRS-CFCRS	**0.632***	**1.195***	**1.524***	**4.225***	**6.824***	**8.155***

can see that removing any component would lead to performance degradation. It indicates that all the components in our model are useful to improve the performance of the recommendation task. Among them, performance decreases the most after removing the template-based dialogue realization. It indicates that the template-based dialogue realization is important in our approach, since it can ensure fluency in language and faithfulness in conversation flow without introducing noise to the simulated data, which is beneficial for the improvement of the recommendation ability of CRSs.

4.3 Evaluation on Conversation Task

In this part, we conduct experiments to verify the effectiveness of our model on the conversation task.

Automatic Evaluation. We show the evaluation results of automatic metrics about different methods in Table 4. As we can see, the methods using PLMs (*i.e.,* GPT-2, DialoGPT, BARCOR, and UniCRS) mostly achieve better performance than other methods. Since PLMs have been pre-trained with generative tasks on large-scale corpora, they can quickly adapt to the CRS task and generate diverse responses after fine-tuning. Among these methods, UniCRS mostly achieves the best performance. Since UniCRS is based on DialoGPT, a PLM that has been pre-trained on large-scale dialogue corpora, it is more capable of generating responses. It also performs semantic fusion and prompt pre-training to inject task-specific knowledge into DialoGPT, helping generate more informative responses.

Besides, we can see that the two data augmentation methods also improve the performance of the three CRS models. Despite the fact that the improvement is not stable, it can demonstrate that CRS models are hungry for training data.

Finally, our model also consistently boosts the performance of the three CRS models, and significantly outperforms the two data augmentation methods. It further indicates the effectiveness and generality of our framework among different CRS methods. Besides, we can see that with the help of our approach, the model KBRD can even outperform PLM-based methods on the ReDial dataset. It shows that our proposed data augmentation approach suits well with KBRD, and can inspire its potential to generate high-quality responses.

Table 5: Human evaluation results about the conversation task on the ReDial dataset.

Models	Fluency	Informativeness
KBRD	0.91	0.86
KBRD-EDA	1.12	1.04
KBRD-mixup	1.05	0.93
KBRD-CFCRS	**1.27**	**1.09**
BARCOR	1.23	1.14
BARCOR-EDA	1.29	1.22
BARCOR-mixup	1.36	1.30
BARCOR-CFCRS	**1.47**	**1.38**
UniCRS	1.41	1.33
UniCRS-EDA	1.57	1.49
UniCRS-mixup	1.48	1.41
UniCRS-CFCRS	**1.69**	**1.60**

Human Evaluation. To provide a more qualified evaluation of the conversation task, we conduct the human evaluation following previous work [52]. We select KBRD, BARCOR, and UniCRS as the backbone, and implement our approach on them. We invite three annotators to evaluate the *fluency* and *informativeness* of the generated responses from examples from these models, and present the results on the ReDial dataset in Table 5. The average Cohen's kappa between any two annotators is 0.89, which indicates good agreement.

First, we can also see a similar tendency to the automatic metrics: UniCRS > BARCOR > KBRD. It indicates the effectiveness of UniCRS which incorporates DialoGPT and knowledge-enhanced prompts. Besides, the two data augmentation methods can consistently improve the quality of the generated response, but their performance order is not stable. A possible reason is that they rely on heuristic rules for augmentation without considering the target model, which may produce useless examples for specific CRS models. Although the augmented examples may contain noise, they are still able to alleviate the data-hungry problem of CRS models. Finally, our approach can consistently outperform these baseline models. It further demonstrates the effectiveness of our framework, which can augment more high-quality examples to improve the training of CRS models, helping them generate fluent and informative responses.

4.4 Performance Comparison w.r.t. Different Amount of Training Data

In real-world applications, the data scarcity problem can be more serious and greatly constrain performance. Since our approach can

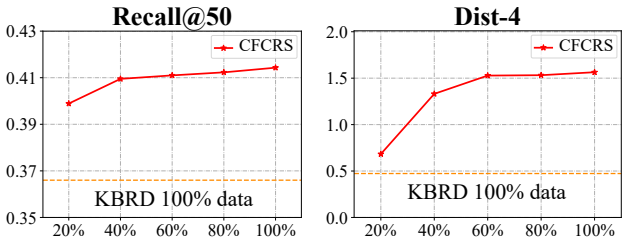

Figure 2: Performance comparison w.r.t. different amounts of training data on the ReDial dataset. We implement our framework on KBRD.

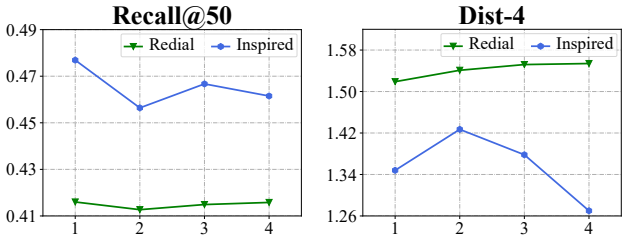

Figure 3: Performance comparison w.r.t. different ratios of augmented examples on ReDial and Inspired dataset. We implement our approach on KBRD.

augment high-quality examples, they can alleviate this problem to some extent. To validate this, we simulate a data scarcity scenario by sampling different proportions of the training data, *i.e.,* 20%, 40%, 60%, 80%, and 100%. We implement our approach on KBRD and report the results of the recommendation and conversation tasks on the ReDial dataset.

Figure 2 shows the evaluation results in different data scarcity settings. As we can see, with just 20% training data, our approach can still achieve comparable performance with KBRD that is trained using 100% data. It indicates that our approach can augment high-quality conversations, greatly alleviating the data scarcity problem. Besides, with less available training data, the performance of our approach is relatively stable. It also shows the potential of our approach to dealing with the cold-start scenario in real-world applications.

4.5 Hyper-parameters Analysis

In our framework, there are two major hyper-parameters to tune: the ratios of augmented examples for each instance and the weights of the L2-norm loss λ during the adversarial training. Here, we investigate the effect of each hyper-parameter on our approach. We conduct the analysis experiments on the recommendation and conversation tasks on the ReDial and INSPIRED datasets. We implement our approach on KBRD and report the results for the two hyper-parameters in Figure 3 and Figure 4, respectively.

First, we can see that for the ReDial dataset, the performance is stable when tuning the two hyper-parameters. It indicates that our approach is not too sensitive to the two hyper-parameters on this dataset. Whereas, too large or too small weights of the L2-norm loss

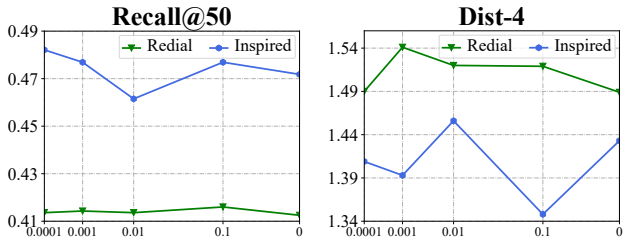

Figure 4: Performance comparison w.r.t. different weights of the L2-norm loss λ on ReDial and Inspired dataset. We implement our approach on KBRD.

λ would cause performance degradation, since too large λ might punish the noise vectors too much, while too small one may bring too much noise. Second, for the INSPIRED dataset, we can see the performance is not stable. A possible reason is that INSPIRED owns very few conversations in the training set, which may hurt the robustness of CRS models. Besides, on the INSPIRED dataset, the best points of the two hyper-parameters in the recommendation and conversation tasks are different. The reason may be that the two tasks focus on different goals, which may lead to conflict in their best hyper-parameter settings.

5 CONCLUSION

In this paper, we proposed a counterfactual data simulation approach, named **CFCRS**, for alleviating the issue of data scarcity in CRSs. We developed our approach under the framework of *counterfactual data augmentation*. Specially, in our approach, we characterized the conversation flow and user preference via the entities mentioned in the conversation. Our approach gradually augmented the user preference from a real dialogue without interfering with the entire conversation flow. Such an augmentation strategy was well learned by an adversarial training method with a curriculum schedule. As a key component, we designed a multi-stage recommendation dialogue simulator based on a flow language model, which can generate *reasonable, coherent* conversation flows for dialogue realization. Extensive experiments have shown that our approach can consistently boost the performance of several competitive CRSs, and outperform other data augmentation methods.

Currently, our approach adopts a multi-stage simulation method to generate CRS data. For future work, we will investigate more unified and simplified approaches for data augmentation, such as the utilization of large language models [38, 49].

ACKNOWLEDGMENTS

We are thankful to Jinhao Jiang for his supportive work. This work was partially supported by National Natural Science Foundation of China under Grant No. 62222215, Beijing Natural Science Foundation under Grant No. 4222027, Beijing Outstanding Young Scientist Program under Grant No. BJJWZYJH012019100020098, the Outstanding Innovative Talents Cultivation Funded Programs 2022 of Renmin University of China. Xin Zhao is the corresponding author.

REFERENCES

[1] Hongshen Chen, Xiaorui Liu, Dawei Yin, and Jiliang Tang. 2017. A survey on dialogue systems: Recent advances and new frontiers. *Acm Sigkdd Explorations Newsletter* 19, 2 (2017), 25–35.

[2] Qibin Chen, Junyang Lin, Yichang Zhang, Ming Ding, Yukuo Cen, Hongxia Yang, and Jie Tang. 2019. Towards Knowledge-Based Recommender Dialog System. In *EMNLP*. 1803–1813.

[3] Konstantina Christakopoulou, Filip Radlinski, and Katja Hofmann. 2016. Towards conversational recommender systems. In *KDD*. 815–824.

[4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *NAACL*. 4171–4186.

[5] Steven Y Feng, Varun Gangal, Jason Wei, Sarath Chandar, Soroush Vosoughi, Teruko Mitamura, and Eduard Hovy. 2021. A survey of data augmentation approaches for NLP. *arXiv preprint arXiv:2105.03075* (2021).

[6] Chongming Gao, Wenqiang Lei, Xiangnan He, Maarten de Rijke, and Tat-Seng Chua. 2021. Advances and challenges in conversational recommender systems: A survey. *AI Open* 2 (2021), 100–126.

[7] Jianfeng Gao, Michel Galley, and Lihong Li. 2018. Neural approaches to conversational AI. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 1371–1374.

[8] Yash Goyal, Ziyan Wu, Jan Ernst, Dhruv Batra, Devi Parikh, and Stefan Lee. 2019. Counterfactual visual explanations. In *International Conference on Machine Learning*. PMLR, 2376–2384.

[9] Jiawei Han, Hong Cheng, Dong Xin, and Xifeng Yan. 2007. Frequent pattern mining: current status and future directions. *Data mining and knowledge discovery* 15, 1 (2007), 55–86.

[10] Shirley Anugrah Hayati, Dongyeop Kang, Qingxiaoyang Zhu, Weiyan Shi, and Zhou Yu. 2020. INSPIRED: Toward Sociable Recommendation Dialog Systems. In *EMNLP*. 8142–8152.

[11] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*. 173–182.

[12] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2016. Session-based Recommendations with Recurrent Neural Networks. In *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.). http://arxiv.org/abs/1511.06939

[13] Yutai Hou, Yijia Liu, Wanxiang Che, and Ting Liu. 2018. Sequence-to-Sequence Data Augmentation for Dialogue Language Understanding. In *Proceedings of the 27th International Conference on Computational Linguistics*. 1234–1245.

[14] Dietmar Jannach, Ahtsham Manzoor, Wanling Cai, and Li Chen. 2021. A survey on conversational recommender systems. *ACM Computing Surveys (CSUR)* 54, 5 (2021), 1–36.

[15] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *2018 IEEE international conference on data mining (ICDM)*. IEEE, 197–206.

[16] Ashutosh Kumar, Satwik Bhattamishra, Manik Bhandari, and Partha Talukdar. 2019. Submodular optimization-based diverse paraphrasing and its effectiveness in data augmentation. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. 3609–3619.

[17] Wenqiang Lei, Xiangnan He, Yisong Miao, Qingyun Wu, Richang Hong, Min-Yen Kan, and Tat-Seng Chua. 2020. Estimation-action-reflection: Towards deep interaction between conversational and recommender systems. In *WSDM*.

[18] Raymond Li, Samira Ebrahimi Kahou, Hannes Schulz, Vincent Michalski, Laurent Charlin, and Chris Pal. 2018. Towards deep conversational recommendations. *NeurIPS* 31 (2018).

[19] Shuokai Li, Ruobing Xie, Yongchun Zhu, Xiang Ao, Fuzhen Zhuang, and Qing He. 2022. User-centric conversational recommendation with multi-aspect user modeling. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 223–233.

[20] Qi Liu, Matt Kusner, and Phil Blunsom. 2021. Counterfactual data augmentation for neural machine translation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 187–197.

[21] Zhiwei Liu, Ziwei Fan, Yu Wang, and Philip S Yu. 2021. Augmenting sequential recommendation with pseudo-prior items via reversely pre-training transformer. In *Proceedings of the 44th international ACM SIGIR conference on Research and development in information retrieval*. 1608–1612.

[22] Zeming Liu, Haifeng Wang, Zheng-Yu Niu, Hua Wu, Wanxiang Che, and Ting Liu. 2020. Towards Conversational Recommendation over Multi-Type Dialogs. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 1036–1049.

[23] Ilya Loshchilov and Frank Hutter. 2018. Decoupled Weight Decay Regularization. In *ICLR*.

[24] Yu Lu, Junwei Bao, Yan Song, Zichen Ma, Shuguang Cui, Youzheng Wu, and Xiaodong He. 2021. RevCore: Review-Augmented Conversational Recommendation. In *ACL Findings*. 1161–1173.

[25] S Chandra Mouli, Yangze Zhou, and Bruno Ribeiro. [n. d.]. Bias Challenges in Counterfactual Data Augmentation. In *UAI 2022 Workshop on Causal Representation Learning*.

[26] Jiao Ou, Jinchao Zhang, Yang Feng, and Jie Zhou. 2022. Counterfactual Data Augmentation via Perspective Transition for Open-Domain Dialogues. *arXiv preprint arXiv:2210.16838* (2022).

[27] Silviu Pitis, Elliot Creager, and Animesh Garg. 2020. Counterfactual data augmentation using locally factored dynamics. *Advances in Neural Information Processing Systems* 33 (2020), 3976–3990.

[28] Ruihong Qiu, Zi Huang, and Hongzhi Yin. 2021. Memory augmented multi-instance contrastive predictive coding for sequential recommendation. In *2021 IEEE International Conference on Data Mining (ICDM)*. IEEE, 519–528.

[29] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog* 1, 8 (2019), 9.

[30] Zhaochun Ren, Zhi Tian, Dongdong Li, Pengjie Ren, Liu Yang, Xin Xin, Huasheng Liang, Maarten de Rijke, and Zhumin Chen. 2022. Variational Reasoning about User Preferences for Conversational Recommendation. In *SIGIR*.

[31] Marco Túlio Ribeiro, Sameer Singh, and Carlos Guestrin. 2018. Semantically Equivalent Adversarial Rules for Debugging NLP models. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers*. 856–865. https://doi.org/10.18653/v1/P18-1079

[32] Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In *European semantic web conference*. Springer, 593–607.

[33] Chenzhan Shang, Yupeng Hou, Wayne Xin Zhao, Yaliang Li, and Jing Zhang. 2023. Multi-grained Hypergraph Interest Modeling for Conversational Recommendation. *arXiv preprint arXiv:2305.04798* (2023).

[34] Connor Shorten and Taghi M Khoshgoftaar. 2019. A survey on image data augmentation for deep learning. *Journal of big data* 6, 1 (2019), 1–48.

[35] Sandeep Subramanian, Adam Trischler, Yoshua Bengio, and Christopher J Pal. 2018. Learning General Purpose Distributed Sentence Representations via Large Scale Multi-task Learning. In *ICLR*.

[36] Lingzhi Wang, Huang Hu, Lei Sha, Can Xu, Daxin Jiang, and Kam-Fai Wong. 2022. RecInDial: A Unified Framework for Conversational Recommendation with Pretrained Language Models. In *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing*. 489–500.

[37] Ting-Chun Wang, Shang-Yu Su, and Yun-Nung Chen. 2022. BARCOR: Towards A Unified Framework for Conversational Recommendation Systems. *arXiv preprint arXiv:2203.14257* (2022).

[38] Xiaolei Wang, Xinyu Tang, Wayne Xin Zhao, Jingyuan Wang, and Ji-Rong Wen. 2023. Rethinking the Evaluation for Conversational Recommendation in the Era of Large Language Models. *arXiv preprint arXiv:2305.13112* (2023).

[39] Xiaolei Wang, Kun Zhou, Ji-Rong Wen, and Wayne Xin Zhao. 2022. Towards Unified Conversational Recommender Systems via Knowledge-Enhanced Prompt Learning. *arXiv preprint arXiv:2206.09363* (2022).

[40] Yicheng Wang and Mohit Bansal. 2018. Robust Machine Comprehension Models via Adversarial Training. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT, New Orleans, Louisiana, USA, June 1-6, 2018, Volume 2 (Short Papers)*. 575–581. https://doi.org/10.18653/v1/n18-2091

[41] Zhenlei Wang, Jingsen Zhang, Hongteng Xu, Xu Chen, Yongfeng Zhang, Wayne Xin Zhao, and Ji-Rong Wen. 2021. Counterfactual data-augmented sequential recommendation. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*. 347–356.

[42] Jason W. Wei and Kai Zou. 2019. EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019*. 6381–6387. https://doi.org/10.18653/v1/D19-1670

[43] Qingsong Wen, Liang Sun, Fan Yang, Xiaomin Song, Jingkun Gao, Xue Wang, and Huan Xu. 2020. Time series data augmentation for deep learning: A survey. *arXiv preprint arXiv:2002.12478* (2020).

[44] Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Reinforcement learning* (1992), 5–32.

[45] Qizhe Xie, Zihang Dai, Eduard H. Hovy, Thang Luong, and Quoc Le. 2020. Unsupervised Data Augmentation for Consistency Training. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*. https://proceedings.neurips.cc/paper/2020/hash/44feb0096faa8326192570788b38c1d1-Abstract.html

[46] Hongyi Zhang, Moustapha Cisse, Yann N Dauphin, and David Lopez-Paz. 2017. mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412*

(2017).

[47] Xiaoyu Zhang, Xin Xin, Dongdong Li, Wenxuan Liu, Pengjie Ren, Zhumin Chen, Jun Ma, and Zhaochun Ren. 2022. Variational Reasoning over Incomplete Knowledge Graphs for Conversational Recommendation. *arXiv preprint arXiv:2212.11868* (2022).

[48] Yizhe Zhang, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, and William B Dolan. 2020. DIALOGPT: Large-Scale Generative Pre-training for Conversational Response Generation. In *ACL*. 270–278.

[49] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al. 2023. A survey of large language models. *arXiv preprint arXiv:2303.18223* (2023).

[50] Jinfeng Zhou, Bo Wang, Zhitong Yang, Dongming Zhao, Kun Huang, Ruifang He, and Yuexian Hou. 2022. CR-GIS: Improving Conversational Recommendation via Goal-aware Interest Sequence Modeling. In *Proceedings of the 29th International Conference on Computational Linguistics*. 400–411.

[51] Kun Zhou, Xiaolei Wang, Yuanhang Zhou, Chenzhan Shang, Yuan Cheng, Wayne Xin Zhao, Yaliang Li, and Ji-Rong Wen. 2021. CRSLab: An Open-Source Toolkit for Building Conversational Recommender System. In *ACL: System Demonstrations*. 185–193.

[52] Kun Zhou, Wayne Xin Zhao, Shuqing Bian, Yuanhang Zhou, Ji-Rong Wen, and Jingsong Yu. 2020. Improving conversational recommender systems via knowledge graph based semantic fusion. In *KDD*. 1006–1014.

[53] Kun Zhou, Wayne Xin Zhao, Hui Wang, Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, and Ji-Rong Wen. 2020. Leveraging historical interaction data for improving conversational recommender system. In *CIKM*. 2349–2352.

[54] Kun Zhou, Yuanhang Zhou, Wayne Xin Zhao, Xiaoke Wang, and Ji-Rong Wen. 2020. Towards Topic-Guided Conversational Recommender System. In *Proceedings of the 28th International Conference on Computational Linguistics*. 4128–4139.

[55] Yuanhang Zhou, Kun Zhou, Wayne Xin Zhao, Cheng Wang, Peng Jiang, and He Hu. 2022. C^2-CRS: Coarse-to-Fine Contrastive Learning for Conversational Recommender System. In *WSDM 2022*. ACM, 1488–1496.

[56] Ran Zmigrod, Sabrina J Mielke, Hanna Wallach, and Ryan Cotterell. 2019. Counterfactual Data Augmentation for Mitigating Gender Stereotypes in Languages with Rich Morphology. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 1651–1661.

[57] Jie Zou, Evangelos Kanoulas, Pengjie Ren, Zhaochun Ren, Aixin Sun, and Cheng Long. 2022. Improving Conversational Recommender Systems via Transformer-based Sequential Modelling. In *SIGIR*. 2319–2324.