

Persuading across Diverse Domains: A Dataset and Persuasion Large Language Model

Chuhao Jin^{1,2*}, Kening Ren*, Lingzhen Kong^{1,2}, Xiting Wang^{1,2†}, Ruihua Song^{1,2†},
Huan Chen³

¹Gaoling School of Artificial Intelligence, Renmin University of China, Beijing, China

²Beijing Key Laboratory of Big Data Management and Analysis Methods

³Meituan, Beijing, China

{jinchuhao, konglingzhen, xitingwang, rsong}@ruc.edu.cn,

rkeningr@gmail.com, chenhuan15@meituan.com

Abstract

Persuasive dialogue requires multi-turn following and planning abilities to achieve the goal of persuading users, which is still challenging even for state-of-the-art large language models (LLMs). Previous works focus on retrieval-based models or generative models in a specific domain due to a lack of data across multiple domains. In this paper, we leverage GPT-4 to create the first multi-domain persuasive dialogue dataset DailyPersuasion. Then we propose a general method named PersuGPT to learn a persuasion model based on LLMs through intent-to-strategy reasoning, which summarizes the intent of user's utterance and reasons next strategy to respond. Moreover, we design a simulation-based preference optimization, which utilizes a learned user model and our model to simulate next turns and estimate their rewards more accurately. Experimental results on two datasets indicate that our proposed method outperforms all baselines in terms of automatic evaluation metric Win-Rate and human evaluation. The code and data are available at <https://persugpt.github.io>.

1 Introduction

Persuasive dialogue is a frequent and important social interaction in daily life. However, it is difficult to mimic such a general ability for a dialogue system because a good persuader requires many skills, such as understanding users' underlying intents (Tian et al., 2020), applying effective persuasive strategies (Jin et al., 2023) according to different situations across domains, and anticipating users' sophisticated reactions (Samad et al., 2022). Previous works have studied retrieval-based models (Jin et al., 2023) or generative models (Chen et al., 2022b) on a small-scale dataset of a specific domain, e.g., PersuasionForGood (Wang et al., 2019) collecting about 1,000 sessions in charity

*Equal Contribution.

†Corresponding authors: Xiting Wang and Ruihua Song.

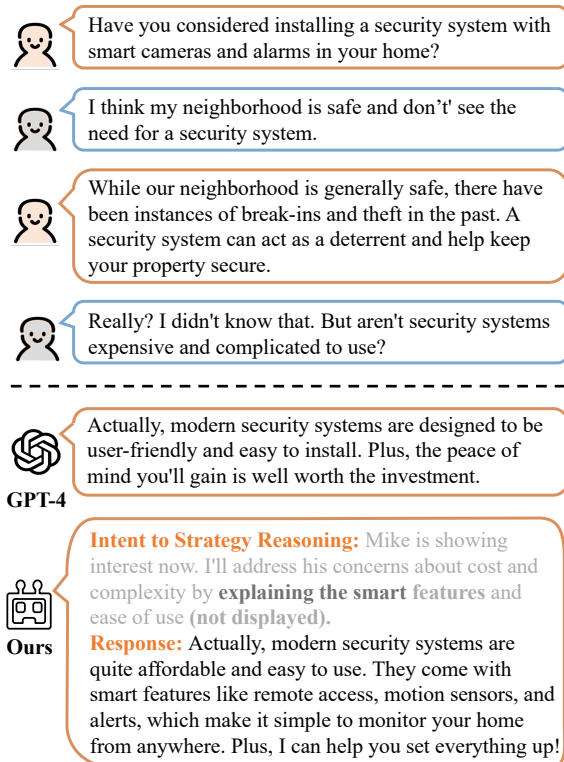


Figure 1: An example of generating persuasive dialogue in the scenario of *persuading Mike to install a security system in home*. The strategy used is in bold.

domain. Recent advances in large language models (LLMs) (Brown et al., 2020; Ouyang et al., 2022) shed light on developing a general persuasion model, but it is still challenging even for state-of-the-art LLMs due to the scarcity of large-scale multi-domain datasets, as well as the need for multi-turn following and planning abilities to achieve the goal of persuading users (Zhou et al., 2024).

In this paper, we aim to address the problem of learning a general persuasion model across domains. First, we construct a multi-domain persuasive dialogue dataset named DailyPersuasion, with the assistance of GPT-4 and a well-designed solution. It comprises 78,000 dialogue sessions spanning 35 domains in daily lives. Second, to enhance multi-turn following and planning ability, we gener-

ate user intents and next strategies while collecting the data and integrate an intent-to-strategy reasoning process into our proposed PersuGPT. A strategy reasoning case is shown in Fig. 1, PersuGPT captures the user’s intent “showing interest”, and then chooses the “explaining the smart features” strategy. Third, we propose using a learned user model and our PersuGPT to simulate the next following turns and evaluate rewards, and then use the composed preference to directly optimize PersuGPT. Extensive experiments on two datasets indicate that our PersuGPT, equipped with 13 billion parameters, outperforms all baselines including GPT-4 in terms of automatic metric Win-Rate and human ratings.

Our main contributions are as follows: (1) To our best knowledge, we construct the first large-scale multi-domain persuasive dialogue dataset **DailyPersuasion**, comprising 78,000 dialogue sessions across 35 domains; (2) We propose a cross-domain persuasion model **PersuGPT**, which integrates the proposed intent-to-strategy reasoning and multi-turn simulation-based preference optimization; and (3) Extensive experiments verify that our model has significant superiority over the state-of-the-art LLMs including GPT-4, and strong generalization abilities across domains.

2 Related Work

2.1 Persuasive Dialogue Datasets and Systems

Recent years have witnessed growing interest in persuasive dialogue, although it is difficult to obtain large-scale persuasion data from humans. Previous works try to collect public persuasion datasets through crowd-sourcing (Wang et al., 2019), engaging interaction between humans and agents in games (Bakhtin et al., 2022) or a social interactive simulation environment (Zhou et al., 2024), but the scale is small. Jin et al. (2023) collects a large-scale debt risk alert dataset while it is not publicly available due to privacy concerns. Chen et al. (2022a) use external chat datasets to enhance the recommendation system. Despite the scarcity of data, previous works have investigated many factors, such as emotional influence (Samad et al., 2022), social facts (Chen et al., 2022b), strategies (Carlisle et al., 2018; Tian et al., 2020; Jin et al., 2023) and intents (Cheng et al., 2023b,c), to build an effective persuasive dialogue system. Different from the prior studies, our approach is not restricted to a specific persuasion domain or limited predefined strategies.

2.2 Cross Domain Adaption of LLMs

Previous works adapt LLMs to new domains through the following ways: 1) Instruction-Tuning, which fine-tunes LLMs on specific datasets, ranging from reformatting existing datasets (Sanh et al., 2022; Wei et al., 2022a), using daily chat data (Ouyang et al., 2022; Köpf et al., 2023), to generating synthetic data (Wang et al., 2022; Xu et al., 2023), and thus generalizes them to new tasks or domains (Lou et al., 2023; Zhao et al., 2023; Yang et al., 2024; Cheng et al., 2023a); 2) Alignment-Tuning, which adjusts LLMs to ensure their behaviors align with human expectations (Zhao et al., 2023; Ouyang et al., 2022) via using LLMs to construct alignment datasets (Sun et al., 2023b; Liu et al., 2022b), reinforcement learning method (Chen et al., 2023; Wang et al., 2021), such as based on human feedback (RLHF) (Ouyang et al., 2022), or supervised learning for model alignment (Lu et al., 2022; Guo et al., 2023), such as CoH (Liu et al., 2023), and DPO (Rafailov et al., 2023); 3) Prompt Engineering, which optimizes the interaction between humans and LLMs via better designing prompts (Liu and Chilton, 2022; White et al., 2023) automatically (Shin et al., 2020; Deng et al., 2022) or manually (Liu et al., 2022a; Rubin et al., 2022), in-context learning (Dong et al., 2023; Brown et al., 2020), and chain of thought (CoT) (Wei et al., 2022b). These kinds of methods enable LLMs to handle new domains and do better reasoning, inspiring us to explore their application in addressing our cross-domain persuasion dialogue problem.

3 Constructing DailyPersuasion Dataset

To ensure the diversity and naturalness of persuasive conversations, as shown in Fig. 2, we adopt a three-step data collection framework to build our DailyPersuasion dataset: $\{(C_i, S_i, D_i)\}_{i=1}^N$, where C_i is a persuasion scenario containing a background and a goal, S_i is a strategy set related to the scenario, and D_i is a session of dialogues corresponded to the scenario and strategies. We evaluate samples to ensure quality and diversity at each step.

3.1 Persuasion Scenario Generation

Although LLMs have shown impressive creativity, when we prompt GPT-4 to generate new persuasion scenarios from examples like Self-Instruct (Wang et al., 2022), we observe that most of the expanded scenarios cover limited common topics such as “fit-

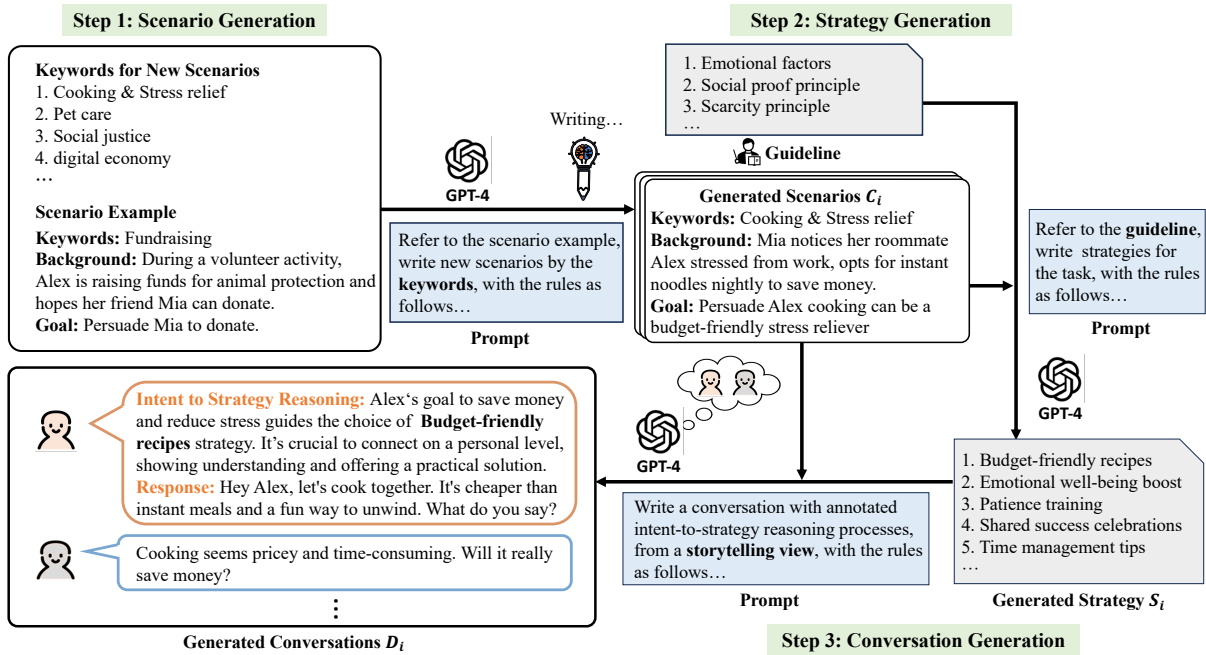


Figure 2: Overview of our solution to collect our multi-domain persuasion dialogue dataset DailyPersuasion.

ness” or “health management.” To address this issue, we propose using topic keywords to induce generating diverse persuasion scenarios. First, we manually identified 128 seed keywords related to persuasion from the MOSS dataset (Sun et al., 2023a), such as “ecological protection” and “product sales”. Then we expand them to 580 keywords with the assistance of GPT-4. To further enhance diversity, we also combine any two keywords as a new compound topic keyword, such as “education & medical”. Next, as shown in Fig. 2, we provide a keyword-to-scenario example in context and prompt GPT-4 to generate persuasion scenarios from the given keywords, e.g., generating “Mia notices her roommate Alex stressed from work, opts for instant noodles nightly to save money” as a background and “Persuade Alex cooking can be a budget-friendly stress reliever” as a goal from the keyword “cooking & stress relief”. Details of prompts and keywords list are shown in Fig. 9 and Tab. 10 of the Appendix.

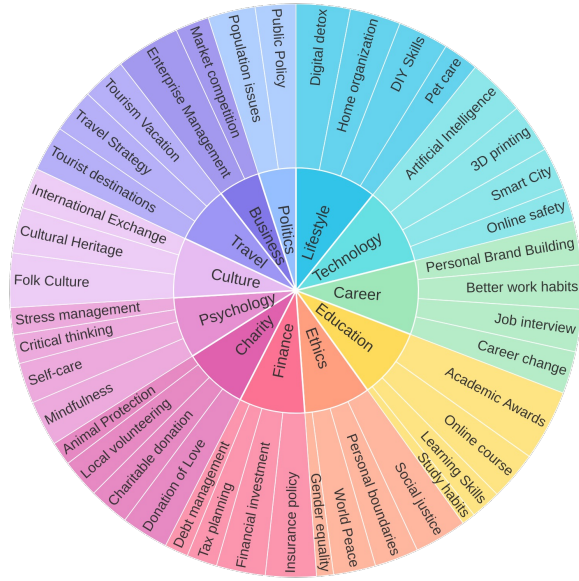
Analysis. DailyPersuasion consists of 13,000 scenarios, which can be grouped into about 35 domains. We show some large domains and selected scenarios in Fig. 3 (a). To measure the diversity of scenarios, we follow the method proposed in Wang et al. (2022) and randomly sample 5,000 scenarios created by our method and by Self-Instruct to calculate the maximum ROUGE-L between each scenario with others, respectively. According to their histograms shown in Fig. 3 (b), the ROUGE-L

scores of our method hover around 0.45 whereas those of Self-Instruct method around 0.6, indicating that our method yields more diverse scenarios. To measure the quality of our scenarios, we randomly select 100 scenarios and hire annotators to evaluate whether a scenario is aligned with the given keyword and whether it is reasonable in common sense. Results indicate that 89% of our scenarios are well aligned and 81% are reasonable. The distribution of scenarios is shown in Fig. 12 of the Appendix.

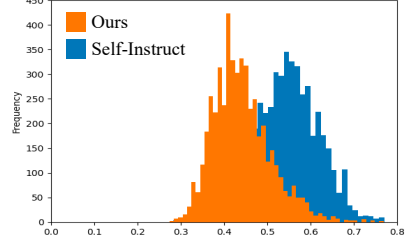
3.2 Persuasion Strategy Generation

We expect strategy sets to reflect a deep comprehension of corresponding scenarios and to be in harmony with the foundational principles of sociology and psychology. Accordingly, we develop a guideline for strategy generation with the inspiration from Cialdini (2001), which offers a variety of principles derived from psychology and sociology, such as “emotional factors” and “social identity principles”. Then, as shown by Step 2 in Fig. 2, we prompt GPT-4 to reference this guideline while generating corresponding strategies. Details are shown in Fig. 10 and Tab. 4 of Appendix.

Analysis. DailyPersuasion consists of 229,598 strategies, which has a long-tail distribution. 214 strategies appear more than 200 times, 7,086 strategies have frequencies between 200 and 10, and 222,298 strategies fewer than 10 times. A word cloud of the most frequent strategies is presented in Fig. 3 (c). The prevalent strategies tend to be



(a). Distribution of Persuasion Scenarios



(b). ROUGE-L with the Most Similar Scenario



(c). Word Cloud of High Frequency Strategies

Figure 3: Some analyses of our DailyPersuasion dataset. In (b), lower Rouge-L scores mean higher diversity.

generic, such as “sharing personal experiences”. In contrast, low-frequency strategies tend to be more scenario-specific, e.g., “highlight Python demand”. It is necessary to generate exclusive strategy sets for different domains. We also display the word clouds of strategies associated with different domains, as shown in Fig. 13 of the Appendix, revealing that each domain tends to have unique strategies.

3.3 Conversation Generation

Previous works collect dialogue contents through crowdsourcing, such as PersuasionForGood (Wang et al., 2019), recruiting human participants, persuading them to donate, and tracing their actions. But this is limited to a few domains, such as donations, as anyone can be a donor. Collecting dialogue for various scenarios is challenging. For example, consider a scenario “a father persuades his rebellious daughter to avoid conflicts with her peers at school.” Recruiting such a father or a rebellious daughter as the participant is difficult. Asking human participants to act as such a father or daughter could also be challenging and introduce biases. Recent research shows that LLMs such as ChatGPT can outperform crowd workers for text-annotation tasks (Gilardi et al., 2023). Motivated by the insight that GPT-4 has access to large-scale data and may have a good understanding of diverse scenarios and users, we take the following steps to help GPT-4 generate dialogues closer to real human utterances.

As shown in Step 3 of Fig. 2, we prompt GPT-

4 to create a dialogue session D_i for the given scenario C_i and strategy set S_i . Specifically, for j -th turn, we create the following:

$$D_{i,j} = [u_{i,j}^-; r_{i,j}; u_{i,j}^+],$$

where $u_{i,j}^-$ is a user’s utterance (regarding the user as “against” party in a debate), $r_{i,j}$ is a paragraph narrating the reasoning process from summarizing the user intent to choosing next strategy for a persuader, $u_{i,j}^+$ is the persuader’s response (regarding the persuader as “for” party in a debate) to $u_{i,j}^-$ and influenced by the $r_{i,j}$.

To generate utterances, an intuitive approach is using two prompts to let GPT-4 play the persuader and user. However, despite instructing GPT-4 with specifically designed prompts to align its utterances with the appropriate role and scenario, the generated responses still tend to be unnatural, lacking immersive quality. For example, when directly prompting GPT-4 to play the role of a rebellious daughter, it may reject with an unnatural statement like “I am an artificial intelligence assistant and do not conflict with others,” or generate statements that lack expected personalized expression. To tackle this issue, we convert the role-playing tasks to third-person narrative tasks, for example, prompting GPT-4 to “Write a story from a third-person view, a dialogue about a father persuades his rebellious daughter to avoid conflicts with her peers at school.” Our experiments find that this can significantly alleviate the unnatural problem,

Table 1: Statistics of two persuasive dialogue datasets.

Statistics	DailyPersu.	Persu.ForGood
Domains	35	1
Scenarios	13,000	1
Dialogue Sessions	76,000	1,017
Strategies per Scenario	50.19	27
Turns per Session	5.08	10.43
Words per $u_{i,j}^+$	39.53	22.96
Words per $u_{i,j}^-$	20.16	15.65

letting GPT-4 simulate various personalized interactions more flexibly. The conversations generated by role-playing and storytelling are shown in Tab. 6 and Tab. 7 of the Appendix. To make GPT-4 better understand the characters in persuasion scenarios, we also prompt GPT-4 with possible user reactions, such as hesitant, impatient, and perfunctory.

We prompt GPT-4 to narrate the entire conversation session from a storytelling perspective, using the third-person point of view, denoted by:

$$D_i = [D_{i,1}; \dots; D_{i,L}]$$

Where L is the turns of a session decided by GPT-4 itself. We prompt GPT-4 each session should not exceed 16 turns. To improve the usability of scenarios and strategies, we generate at most 6 dialogue sessions for each scenario. Details of our prompts are shown in Fig. 11 of the Appendix.

Analysis. DailyPersuasion consists of 76,000 sessions. We randomly sample 100 sessions and ask annotators assess whether a session is composed of natural conversations and whether it follows the given scenario. Results indicate that 96% sessions by our method are natural and follow scenarios, whereas, only 84% sessions by two role-playing GPT-4s are natural. Furthermore, we employ ChatGPT to assess the consistency between 600 randomly sampled pairs of strategy and response. Results show that 99.1% pairs are consistent. A fully generated conversation examples are shown in Tab. 9 of the Appendix.

3.4 Comparing DailyPersuasion with PersuasionForGood

The statistics of our DailyPersuasion dataset and the public PersuasionForGood dataset are presented in Tab. 1. Compared to PersuasionForGood, DailyPersuasion has about 75 times of sessions belonging to 13,000 scenarios across about 35 domains, which is much more diverse. Ours has more strategies too, i.e., 50 vs. 27 per scenario. As PersuasionForGood is created by crowd-sourcing, it

has longer sessions, 10 turns vs. 5 turns per session, while our dataset has longer utterances in terms of the number of words per utterance. To our knowledge, DailyPersuasion is the first large-scale multi-domain persuasive dialogue dataset.

4 Training PersuGPT

Based on our DailyPersuasion dataset, we build a large language model named **PersuGPT**, which possesses high persuasion ability across multiple domains. Given a base large language model, our goal is to train the model so that it can 1) persuade more effectively by explicitly reasoning about the user intent and persuasion strategy; and 2) better anticipate user responses and future rewards through simulated interactions. To this end, we develop a two-phase training procedure described as follows.

4.1 Fine-Tuning with Intent-to-Strategy Reasoning

In this phase, we fine-tune the model so that it learns to effectively reason about the user intent and persuasion strategies before generating responses. This fundamental reasoning capability not only increases the probability of generating more persuasive responses but also facilitates cross-domain transferring, as shown in our experiments (Tab. 2). With our collected dataset DailyPersuasion, which contains reasoning paths from GPT-4, this can be achieved by supervised fine-tuning:

$$\arg \max_{\Theta} p(S_i | C_i; \Theta) \times \prod_{j=1}^L p(r_{i,j}, u_{i,j}^+ | C_i, S_i, D_{i,1:j-1}; \Theta) \quad (1)$$

where Θ is the parameters of the base large language model (LLaMA-2 Chat in this paper). Given a persuasion scenario C_i as the initial input, we optimize the first term in Equation 1 to let PersuGPT explicitly generate a candidate strategy set S_i (e.g., “Clarify project goals; ...; Relate to tight deadlines;”). This allows PersuGPT to adapt to unseen persuasion scenarios. Then, given scenario C_i , the generated strategy set S_i and dialogue history $D_{i,1:j-1}$ as input, we maximize the second term in Equation 1, which allows PersuGPT to learn selecting a proper strategy $r_{i,j}$ and generate utterance $u_{i,j}^+$ accordingly. A training example is shown in Tab. 8 of the Appendix. We find that adding the generation of candidate strategies during fine-tuning enables the model to generate more diverse responses and better adapt to the background and goal of a specific persuasion scenario.

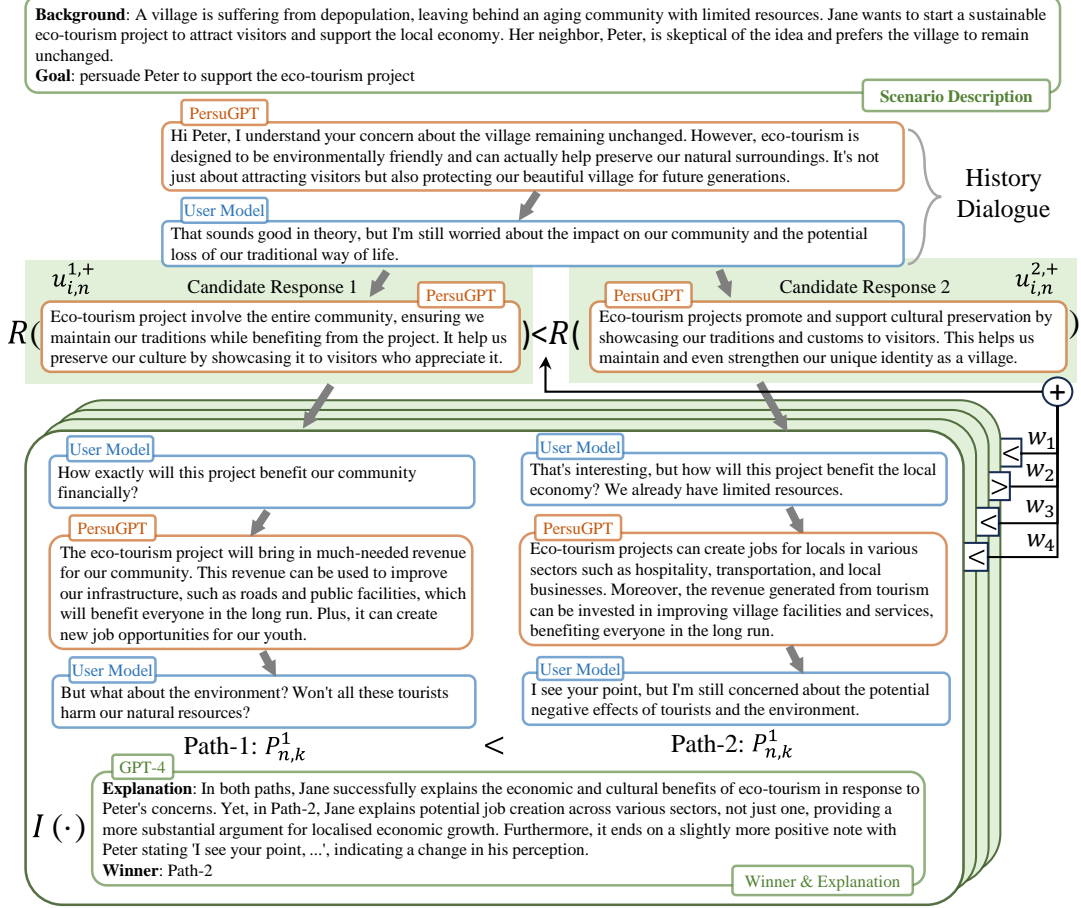


Figure 4: Illustration of our simulation-based preference optimization. Given a scenario, a history dialogue, and two candidate responses ($u_{i,n}^{1,+}$ and $u_{i,n}^{2,+}$), we estimate the long-term reward of the two responses by simulating subsequent conversation paths. Their comparison results of multiple turns are aggregated for preference optimization.

4.2 Simulation-based Preference Optimization

The goal of the second phase is to better align our model with user preferences according to future feedback. This mimics experienced human persuaders, who refine their skills through rigorous practice and continuously honing their utterances based on user feedback. In multi-turn persuasion conversion, aligning a model is uniquely challenging due to the long-term impact of a persuasion strategy and the difficulty of anticipating user feedback in later turns.

To solve this issue, we propose a simulation-based preference optimization method, as shown in Fig. 4. Our method contains the following steps:

1). *User model construction.* We finetune the chat version of the LLaMA model (13 billion parameters) as our user model. We sequentially concatenate the persuasion scenario C_i , the history dialogues $D_{i,1:j-1}$, and the persuader's utterances $u_{i,j}^+$ as input, then we train the model to generate

the user's responses $u_{i,j}^-$.

2). *Candidate response generation.* Given a scenario and a history dialogue of previous $n-1$ turns, we first generate two candidate responses by sampling from our model, PersuGPT. The candidate responses are denoted as $u_{i,n}^{1,+}$ and $u_{i,n}^{2,+}$.

3). *Path simulation.* To estimate the long-term impact of the two candidate responses, we simulate future conversation paths. We let the user model and PersuGPT engage in k turns of follow-up interaction to simulate conversation paths $P_{n,k}^1$ and $P_{n,k}^2$, each starting from one candidate response.

4). *Reward estimation.* We then let GPT-4 compare the two paths and return $I(P_{n,j}^1, P_{n,j}^2)$, where I takes a value of 1 or 0, signifying $P_{n,j}^1$ is superior or not. To comprehensively evaluate each path, we prompt GPT-4 to evaluate from multiple dimensions, such as relevance, emotional factors, persuasiveness, etc. Details are shown in Tab. 5 of the Appendix. The path simulation and GPT-4-based estimation process is repeated for k times,

allowing us to estimate the longer-term reward of the candidate paths by aggregating I :

$$R(u_{i,n}^{1,+}) = \sum_{j=1}^k w_j \cdot I(P_{n,j}^1, P_{n,j}^2) \quad (2)$$

The reward for $u_{i,n}^{2,+}$ is calculated by replacing $I(\cdot)$ with $1 - I(\cdot)$. Weight w_j decreases with increasing conversation length to focus more on the initial turns of conversation that play a more critical role in persuading others. To measure how reliable the GPT-4-based reward is, we ask human labelers to annotate 100 simulated path pairs (decide which path in a pair is better). Results show that 85% of the human labels are consistent with GPT-4 decisions. This result reveals a consistency between GPT-4 and human preferences.

4). *Preference optimization.* Given reward r , we further optimize PersuGPT to favor the generation of higher-reward utterances through Direct Preference Optimization (DPO) (Rafailov et al., 2023).

5 Experiments

We conduct experiments to assess our model’s performance on unseen persuasion scenarios and its adaptability to new domains.

5.1 Setup

We introduce baselines and evaluation metrics, with more details provided in Appendix A.





Baselines. We compare our proposed method with baselines of three categories: 1) **Off-the-shelf LLMs.** Upon our DailyPersuasion dataset, we take ChatGPT and GPT-4 as our baselines and incorporate intent-to-strategy reasoning (ISR) into them as stronger baselines as well. We also compare with in-context learning (ICL) methods, by guiding ChatGPT and GPT-4 with a high-quality dialogue example from our DailyPersuasion dataset. 2) **Fine-tuned LLMs.** We take the chat version of LLaMA-2 model with 13 billion parameters as base model and fine-tune it by using dialogues only (denoted by LLaMA-2 ) and by using dialogues incorporated with ISR on DailyPersuasion as our baselines. We also include the zero-shot of the LLaMA-2 model to validate the effectiveness of our proposed dataset. 3) **Domain-Specific Models.** In the charity domain on which PersuasionForGood focuses, we take the fine-tuned conversation generative model Blenderbot (one of the best models before the LLM era) with 400M parameters as a baseline. In addition,

Table 2: Comparing our PersuGPT with baseline methods on DailyPersuasion dataset.  indicates fine-tuning. “ISR” means using Intent-to-Strategy Reasoning. “ICL” means using In-Context Learning. GPT-4+ISR is used to be compared in counting Win-Rate and generate ground-truth for calculating ROUGE-L.

Model	ICL	ISR	Win-Rate	ROUGE-L	Human
LLaMA-2			20.1	17.4	2.35
LLaMA-2 			46.7	24.4	4.19
ChatGPT			42.0	21.8	3.91
GPT-4			46.9	18.8	4.01
ChatGPT	✓		35.8	22.6	-
GPT-4	✓		46.4	18.9	-
LLaMA-2 		✓	47.8	25.6	4.22
ChatGPT		✓	35.5	20.7	3.83
GPT-4		✓	-	-	4.17
Ours		✓	60.4	25.2	4.35

we fine-tune the LLaMA-2 model and integrate ISR based on the human-labeled intents and strategies provided by PersuasionForGood or our predicted intents and strategies as baselines, assessing which is easier to adapt to a new domain.

Evaluation. We apply the following three metrics, two for automatic evaluation and one for human evaluation: 1) **Win-Rate** is the percentage of winning by comparing a model generated utterance with the ground truth utterance generated by GPT-4 with ISR integrated using ChatGPT as a judge (we have a pair judged twice by switching the order to avoid ordering bias); 2) **Rouge-L** score assesses the similarity between the generated and ground truth utterances at the character level, which is usually only a reference for generative models as ground truth utterances are impossible to cover all reasonable generated results; 3) **Human Rating** is average of three annotators ratings on a scale of 1 to 5, with higher rating indicating better quality. Given a context, we shuffle all utterances for comparison and make them anonymous when showing them to annotators. Thus annotators can compare them before giving a fair judgement.

5.2 Results on Unseen Scenarios

As illustrated in Tab. 2, our proposed model performs the best in terms of Win-Rate and human rating. It achieves a Win-Rate of 60.4%, which clearly exceeds the GPT-4 generated responses, meanwhile receiving a human rating of 4.35, which exceeds the 4.17 score from GPT-4 integrated with intent-to-strategy reasoning (ISR). The human rating of LLaMA-2 is only 2.35, while the fine-tuned

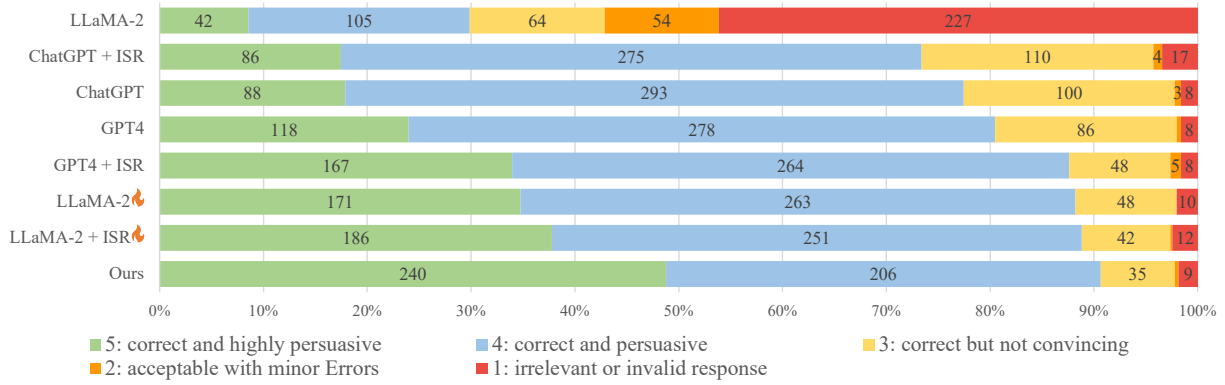


Figure 5: Distribution of human ratings in different levels for all compared methods on DailyPersuasion dataset. Best viewed in color. “ISR” denotes Intent-to-Strategy Reasoning.

Table 3: Comparing our PersuGPT with finetuned baseline methods on PersuasionForGood Dataset. “ISR” is the proposed Intent-to-Strategy Reasoning method, “Auto” means from classifiers trained on our DailyPersuasion dataset, and “Manual” means from PersuasionForGood. “GT” means ground truth from humans.

Model	ISR	Win-Rate	ROUGE-L	Human
Blenderbot		47.1	14.1	3.46
LLaMA-2		46.7	14.6	3.29
LLaMA-2+ISR	Manual	47.7	13.8	3.43
LLaMA-2+ISR	Auto	46.4	13.1	3.28
Human (GT)		-	-	3.45
Ours		52.2	12.8	3.45
Ours+ISR	Manual	53.3	12.4	3.50
Ours+ISR	Auto	55.0	12.7	3.57

LLaMA-2 model reaches 4.19, affirming the effectiveness of our constructed data. Models of GPT-4 and LLaMA-2 🔥 with ISR perform slightly better than those without ISR versions, validating the effectiveness of our proposed intent-to-strategy reasoning. Only the ChatGPT with ISR performs worse than its without ISR version. This may attribute to its inadequate reasoning capabilities, hindering its ability to infer intents and strategies. In contrast, the GPT-4 which has stronger reasoning abilities, can significantly enhance performance when intents and strategies are employed. Compared with the zero-shot, the performance of both LLMs using in-context learning (ICL) decreases, especially for ChatGPT. We found that ChatGPT and GPT-4 sometimes confuse the example with historical conversations. This may be because LLMs can learn from historical conversations, and in our persuasion domain, additional conversation examples may interfere with it.

Fig. 5 gives distribution of ratings for all methods. It indicates that our constructed cross-domain DailyPersuasion dataset is effective in fine-tuning a

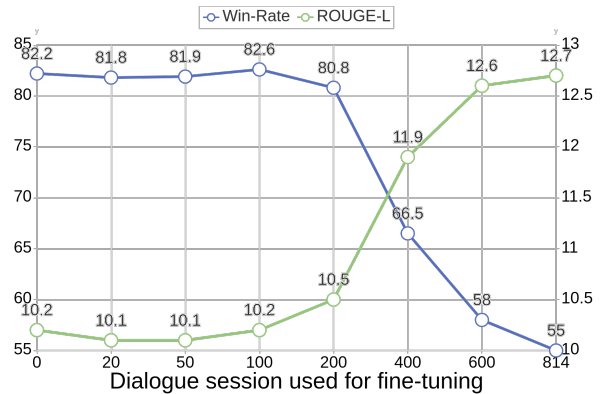


Figure 6: Changes of Win-Rate and Rouge-L scores of our PersuGPT when increasing the amount of fine-tuning data on PersuasionForGood.

better persuasion model, in particular with the help of intent-to-strategy reasoning. Our proposed simulation based preference optimization can further significantly improve the ratio of correct and highly persuasive conversations over LLaMA-2+ISR 🔥.

5.3 Results on a Specific Domain

In real applications, people may target a specific domain. Thus we conduct experiments to compare methods on PersuasionForGood, which focuses on a charity domain while DailyPersuasion has no overlapped data. As shown in Tab. 3, our model after being fine-tuned with our automatically classified intents and strategies (we train two classifiers by using DailyPersuasion data, in which one takes user utterance as input and predict user intent while the other takes next persuader utterance as input and predict its strategy), i.e., Ours+ISR(Auto), outperforms all other generative models in terms of Win-Rate and human rating. If compared with the original collected human responses in ROUGE-L, ours is not as high as Blenderbot. This indicates that using the DailyPersuasion pre-trained model

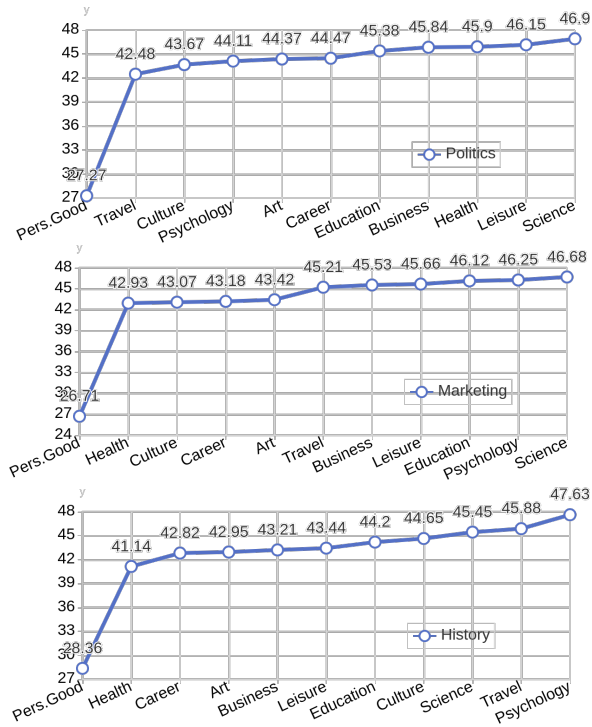


Figure 7: We select 200 scenarios from each of the 10 domains for training ten PersuGPT models and testing them on History, Marketing, and Politics domains. We also test the generalization ability of PersuGPT model trained on PersuasionForGood, as *Pers.Good*. Sort the models in ascending order of Win-Rate.

has its own style, which may be different from the target domain, although the persuasiveness is better from the perspective of user experiences.

We also investigate the impact of fine-tuning data scales on our model performance in PersuasionForGood. Results are shown in Fig. 6. We find that with the data increased, the Rouge-L score rises from 10.2 to 12.7, indicating our model’s growing integration into this specific domain. Conversely, the Win-rate decreases from 82.2% to 55.0%, suggesting that while our model possesses strong persuasive capabilities, adapting it for a specific scenario sacrifices its general persuasiveness, or preference (perhaps in style) of the judge ChatGPT. A case study is shown in Appendix B.

5.4 Generalization across Domains

We investigate the generalization of our proposed method across domains. We randomly select 10 domains from DailyPersuasion and sample 200 scenarios from each to train ten models, and then test them on the target History, Marketing, and Politics domains. As a comparison, we also test the model trained on PersuasionForGood. As shown in Fig. 7, the model trained on domain data of

DailyPersuasion demonstrates significantly higher generalization capabilities than the model trained on PersuasionForGood, indicating the robustness of our data. We observe that the generalization ability of models trained across various domains depends on two key factors: 1) it is related to the intrinsic property of a training domain, e.g., the model trained on the Science domain consistently perform well across three test domains; and 2) it is influenced by the similarity between domains, such as Travel and Culture, performing better when generalizing to a similar domain History.

6 Conclusion

In this paper, we build the first large-scale cross-domain persuasive dialogue dataset named DailyPersuasion, encompassing 76,000 sessions across 35 domains, annotated with user intents and persuader strategies. Then we propose a cross-domain persuasive dialogue model PersuGPT, which incorporates proposed intent-to-strategy reasoning and simulation-based preference optimization. Extensive experiments demonstrate that PersuGPT surpasses GPT-4 with ISR in terms of Win-Rate (60.4%) and human ratings (4.35 vs. 4.17) on DailyPersuasion and all baselines on PersuasionForGood. Our proposed PersuGPT also has good generalization abilities from a training domain to a new test domain. We plan to further optimize PersuGPT through higher-quality automatic interactions and more accurate feedback. Incorporating retrieval-augmented methods for persuasive fact referencing is also within the scope of our future work.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No. 62276268), Beijing Zhidemai Technology Co., Ltd, the Beijing Outstanding Young Scientist Program NO. BJJWZYJH012019100020098, the Public Computing Cloud, Renmin University of China, the Major Innovation & Planning Interdisciplinary Platform for the “Double-First Class” Initiative, Renmin University of China, the Public Policy and Decision-making Research Lab of RUC, the Engineering Research Center of Next-Generation Intelligent Search and Recommendation, Ministry of Education, the fund for building world-class universities (disciplines) of Renmin University of China.

Limitations

Despite our efforts to guide GPT-4 in generating natural dialogues, there may still be inconsistency between our dataset and the distribution of persuasive conversations in the real world, which could limit its effectiveness in certain domains or scenarios. However, given the challenges of collecting high-quality persuasive data in the real world, we believe our dataset remains a valuable asset, particularly for research involving privacy-sensitive or scarce data. Considering the complexity and diversity of human persuasive dialogue, our trained user model may not yet achieve the same level of personality as real human interactions. Nonetheless, as a preliminary exploration of preference optimization based on simulated interactions, this user model is sufficient to test whether user feedback contributes to model refinement.

Ethical Statement

Persuasive dialogue systems serve as a double-edged sword. On one hand, they can be extensively applied in psychological therapy and philanthropic efforts, fostering positive developments within human society. On the other hand, their misuse in potentially harmful scenarios must be strictly regulated. In our study, we filter the keywords used to construct persuasive scenarios, ensuring all generated scenarios are safe and free from bias. We utilize GPT-4, aligned with ethical values, to collect data, hoping to guarantee the gathered data devoid of user privacy breaches and harmful content as GPT-4. We will ask humans to review all scenarios, dialogues, and strategies before releasing DailyPersuasion and further filter inappropriate or risky data out. We will also ask all people or organization that download the dataset to sign strict license to manage the use of our data. It is worth noting that, while our system can be employed across various persuasive domains, it should not be used to directly replace human interaction. All applications of our system should operate under human supervision and regulation, maintaining a balance between leveraging technology for good and ensuring ethical use.

References

A Bakhtin, N Brown, E Dinan, G Farina, C Flaherty, D Fried, A Goff, J Gray, H Hu, AP Jacob, et al. 2022. Human-level play in the game of diplomacy by

combining language models with strategic reasoning. *Science (New York, NY)*, pages eade9097–eade9097.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. [Language models are few-shot learners](#). In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*.

Winston Carlile, Nishant Gurrupadi, Zixuan Ke, and Vincent Ng. 2018. [Give me more feedback: Annotating argument persuasiveness and related attributes in student essays](#). 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*, pages 621–631. Association for Computational Linguistics.

Changyu Chen, Xiting Wang, Yiqiao Jin, Victor Ye Dong, Li Dong, Jie Cao, Yi Liu, and Rui Yan. 2023. Semi-offline reinforcement learning for optimized text generation. In *International Conference on Machine Learning*, pages 5087–5103. PMLR.

Changyu Chen, Xiting Wang, Xiaoyuan Yi, Fangzhao Wu, Xing Xie, and Rui Yan. 2022a. Personalized chit-chat generation for recommendation using external chat corpora. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, pages 2721–2731.

Maximillian Chen, Weiyang Shi, Feifan Yan, Ryan Hou, Jingwen Zhang, Saurav Sahay, and Zhou Yu. 2022b. [Seamlessly integrating factual information and social content with persuasive dialogue](#). 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, AACL/IJCNLP 2022 - Volume 1: Long Papers, Online Only, November 20-23, 2022*, pages 399–413. Association for Computational Linguistics.

Xuxin Cheng, Bowen Cao, Qichen Ye, Zhihong Zhu, Hongxiang Li, and Yuexian Zou. 2023a. MI-Imcl: Mutual learning and large-margin contrastive learning for improving asr robustness in spoken language understanding. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 6492–6505.

Xuxin Cheng, Zhihong Zhu, Bowen Cao, Qichen Ye, and Yuexian Zou. 2023b. Mrrl: Modifying the reference via reinforcement learning for non-autoregressive joint multiple intent detection and slot

- filling. In *The 2023 Conference on Empirical Methods in Natural Language Processing*, pages 10495–10505.
- Xuxin Cheng, Zhihong Zhu, Wanshi Xu, Yaowei Li, Hongxiang Li, and Yuexian Zou. 2023c. Accelerating multiple intent detection and slot filling via targeted knowledge distillation. In *The 2023 Conference on Empirical Methods in Natural Language Processing*.
- Robert B Cialdini. 2001. The science of persuasion. *Scientific American*, 284(2):76–81.
- Mingkai Deng, Jianyu Wang, Cheng-Ping Hsieh, Yihan Wang, Han Guo, Tianmin Shu, Meng Song, Eric P. Xing, and Zhiting Hu. 2022. [Rlprompt: Optimizing discrete text prompts with reinforcement learning](#). In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 3369–3391. Association for Computational Linguistics.
- Qingxiu Dong, Lei Li, Damai Dai, Ce Zheng, Zhiyong Wu, Baobao Chang, Xu Sun, Jingjing Xu, Lei Li, and Zhifang Sui. 2023. [A survey for in-context learning](#). *CoRR*, abs/2301.00234.
- Fabrizio Gilardi, Meysam Alizadeh, and Maël Kubli. 2023. Chatgpt outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, 120(30):e2305016120.
- Geyang Guo, Ranchi Zhao, Tianyi Tang, Wayne Xin Zhao, and Ji-Rong Wen. 2023. [Beyond imitation: Leveraging fine-grained quality signals for alignment](#). *CoRR*, abs/2311.04072.
- Chuhao Jin, Yutao Zhu, Lingzhen Kong, Shijie Li, Xiao Zhang, Ruihua Song, Xu Chen, Huan Chen, Yuchong Sun, Yu Chen, and Jun Xu. 2023. [Joint semantic and strategy matching for persuasive dialogue](#). In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 4187–4197, Singapore. Association for Computational Linguistics.
- Andreas Köpf, Yannic Kilcher, Dimitri von Rütte, Sotiris Anagnostidis, Zhi-Rui Tam, Keith Stevens, Abdullah Barhoum, Nguyen Minh Duc, Oliver Stanley, Richárd Nagyfi, Shahul ES, Sameer Suri, David Glushkov, Arnav Dantuluri, Andrew Maguire, Christoph Schuhmann, Huu Nguyen, and Alexander Mattick. 2023. [Openassistant conversations - democratizing large language model alignment](#). *CoRR*, abs/2304.07327.
- Hao Liu, Carmelo Sferrazza, and Pieter Abbeel. 2023. [Chain of hindsight aligns language models with feedback](#). *CoRR*, abs/2302.02676.
- Jiachang Liu, Dinghan Shen, Yizhe Zhang, Bill Dolan, Lawrence Carin, and Weizhu Chen. 2022a. [What makes good in-context examples for gpt-3?](#) In *Proceedings of Deep Learning Inside Out: The 3rd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures, DeeLIO@ACL 2022, Dublin, Ireland and Online, May 27, 2022*, pages 100–114. Association for Computational Linguistics.
- Ruibo Liu, Chenyan Jia, Ge Zhang, Ziyu Zhuang, Tony X. Liu, and Soroush Vosoughi. 2022b. [Second thoughts are best: Learning to re-align with human values from text edits](#). In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*.
- Vivian Liu and Lydia B. Chilton. 2022. [Design guidelines for prompt engineering text-to-image generative models](#). In *CHI '22: CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 29 April 2022 - 5 May 2022*, pages 384:1–384:23. ACM.
- Renze Lou, Kai Zhang, and Wenpeng Yin. 2023. Is prompt all you need? no. a comprehensive and broader view of instruction learning. *arXiv preprint arXiv:2303.10475*.
- Ximing Lu, Sean Welleck, Jack Hessel, Liwei Jiang, Lianhui Qin, Peter West, Prithviraj Ammanabrolu, and Yejin Choi. 2022. [QUARK: controllable text generation with reinforced unlearning](#). In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*.
- Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F. Christiano, Jan Leike, and Ryan Lowe. 2022. [Training language models to follow instructions with human feedback](#). In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*.
- Rafael Rafailov, Archit Sharma, Eric Mitchell, Stefano Ermon, Christopher D. Manning, and Chelsea Finn. 2023. [Direct preference optimization: Your language model is secretly a reward model](#). *CoRR*, abs/2305.18290.
- Ohad Rubin, Jonathan Herzig, and Jonathan Berant. 2022. [Learning to retrieve prompts for in-context learning](#). In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL 2022, Seattle, WA, United States, July 10-15, 2022*, pages 2655–2671. Association for Computational Linguistics.
- Azlaan Mustafa Samad, Kshitij Mishra, Mauajama Firdaus, and Asif Ekbal. 2022. [Empathetic persuasion: Reinforcing empathy and persuasiveness in dialogue](#)

- systems. In *Findings of the Association for Computational Linguistics: NAACL 2022, Seattle, WA, United States, July 10-15, 2022*, pages 844–856. Association for Computational Linguistics.
- Victor Sanh, Albert Webson, Colin Raffel, Stephen H. Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Arun Raja, Manan Dey, M Saiful Bari, Canwen Xu, Urmish Thakker, Shanya Sharma Sharma, Eliza Szczechla, Taewoon Kim, Gunjan Chhablani, Nihal V. Nayak, Debajyoti Datta, Jonathan Chang, Mike Tian-Jian Jiang, Han Wang, Matteo Manica, Sheng Shen, Zheng Xin Yong, Harshit Pandey, Rachel Bawden, Thomas Wang, Trishala Neeraj, Jos Rozen, Abheesht Sharma, Andrea Santilli, Thibault Févry, Jason Alan Fries, Ryan Teehan, Teven Le Scao, Stella Biderman, Leo Gao, Thomas Wolf, and Alexander M. Rush. 2022. [Multi-task prompted training enables zero-shot task generalization](#). In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net.
- Taylor Shin, Yasaman Razeghi, Robert L. Logan IV, Eric Wallace, and Sameer Singh. 2020. [Autoprompt: Eliciting knowledge from language models with automatically generated prompts](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, pages 4222–4235. Association for Computational Linguistics.
- Tianxiang Sun, Xiaotian Zhang, Zhengfu He, Peng Li, Qinyuan Cheng, Hang Yan, Xiangyang Liu, Yunfan Shao, Qiong Tang, Xingjian Zhao, Ke Chen, Yining Zheng, Zhejian Zhou, Ruixiao Li, Jun Zhan, Yunhua Zhou, Linyang Li, Xiaogui Yang, Lingling Wu, Zhangyue Yin, Xuanjing Huang, and Xipeng Qiu. 2023a. Moss: Training conversational language models from synthetic data.
- Zhiqing Sun, Yikang Shen, Qinrong Zhou, Hongxin Zhang, Zhenfang Chen, David D. Cox, Yiming Yang, and Chuang Gan. 2023b. [Principle-driven self-alignment of language models from scratch with minimal human supervision](#). *CoRR*, abs/2305.03047.
- Youzhi Tian, Weiyan Shi, Chen Li, and Zhou Yu. 2020. [Understanding user resistance strategies in persuasive conversations](#). In *Findings of the Association for Computational Linguistics: EMNLP 2020, Online Event, 16-20 November 2020*, volume EMNLP 2020 of *Findings of ACL*, pages 4794–4798. Association for Computational Linguistics.
- Xiting Wang, Xinwei Gu, Jie Cao, Zihua Zhao, Yulan Yan, Bhuvan Middha, and Xing Xie. 2021. Reinforcing pretrained models for generating attractive text advertisements. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 3697–3707.
- Xuwei Wang, Weiyan Shi, Richard Kim, Yoojung Oh, Sijia Yang, Jingwen Zhang, and Zhou Yu. 2019. [Persuasion for good: Towards a personalized persuasive dialogue system for social good](#). In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pages 5635–5649. Association for Computational Linguistics.
- Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A Smith, Daniel Khashabi, and Han-naneh Hajishirzi. 2022. [Self-instruct: Aligning language model with self generated instructions](#). *arXiv preprint arXiv:2212.10560*.
- Jason Wei, Maarten Bosma, Vincent Y. Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, and Quoc V. Le. 2022a. [Finetuned language models are zero-shot learners](#). In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net.
- Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H. Chi, Quoc V. Le, and Denny Zhou. 2022b. [Chain-of-thought prompting elicits reasoning in large language models](#). In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*.
- Jules White, Quchen Fu, Sam Hays, Michael Sandborn, Carlos Olea, Henry Gilbert, Ashraf Elnashar, Jesse Spencer-Smith, and Douglas C. Schmidt. 2023. [A prompt pattern catalog to enhance prompt engineering with chatgpt](#). *CoRR*, abs/2302.11382.
- Can Xu, Qingfeng Sun, Kai Zheng, Xiubo Geng, Pu Zhao, Jiazhan Feng, Chongyang Tao, and Daxin Jiang. 2023. [Wizardlm: Empowering large language models to follow complex instructions](#). *CoRR*, abs/2304.12244.
- Weikai Yang, Mengchen Liu, Zheng Wang, and Shixia Liu. 2024. Foundation models meet visualizations: Challenges and opportunities. *Computational Visual Media*, pages 1–26.
- 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*.
- Xuhui Zhou, Hao Zhu, Leena Mathur, Ruohong Zhang, Haofei Yu, Zhengyang Qi, Louis-Philippe Morency, Yonatan Bisk, Daniel Fried, Graham Neubig, and Maarten Sap. 2024. [SOTOPIA: Interactive evaluation for social intelligence in language agents](#). In *The Twelfth International Conference on Learning Representations*.

A Experiment Setup

A.1 Implementation Details

We split the DailyPersuasion dataset into three subsets: 10,000 scenarios for fine-tuning, 2,000

for preference optimization, and 1,000 for testing. Our model is developing from the LLaMA-2 Chat model with 13 billion parameters. We train the model for one epoch in the fine-tuning phase. For preference optimization, we randomly sample 2,000 historical dialogues of varying lengths, which are also trained in one epoch. We assess the model’s domain adaptability on the public PersuasionForGood dataset (Wang et al., 2019). To validate the effectiveness of the intent-to-strategy reasoning on this dataset, we align the provided strategy and intent labels into a uniform format. In addition, we employ a predictor trained on the DailyPersuasion training set for the intent-to-strategy reasoning to annotate the PersuasionForGood dataset, ensuring consistency in data distribution with DailyPersuasion. Our training and evaluation on PersuasionForGood utilize 814 and 203 sessions, respectively.

A.2 Evaluation Dataset Preparation

We evaluate models on the proposed DailyPersuasion dataset, assessing its performance across 1,000 unseen persuasion scenarios. We also compare models on 203 sessions from the PersuasionForGood dataset, to gauge its domain-specific performance. We sample dialogue contexts with random turns from the test set and generate subsequent responses using various methods. Then we evaluate the quality of the generated responses. We sample 492 and 120 context-response pairs from the test set of DailyPersuasion and PersuasionForGood for human evaluation.

B Case Study

By examining cases from PersuasionForGood, we explore why PersuGPT outperforms humans in persuasion. As illustrated in Fig. 8, when a user shows interest in the organization, PersuGPT accurately captures her positive attitude through reasoning. It then guides the conversation toward persuasive responses, such as employing the word “amazing” to enhance emotional impact and using vivid storytelling to engage the user. In contrast, humans often struggle to balance semantic coherence with persuasion or fail to employ persuasion strategies efficiently. This confirms that PersuGPT can effectively use strategies and integrate them into responses more appropriately.

C Prompts for Data Collection

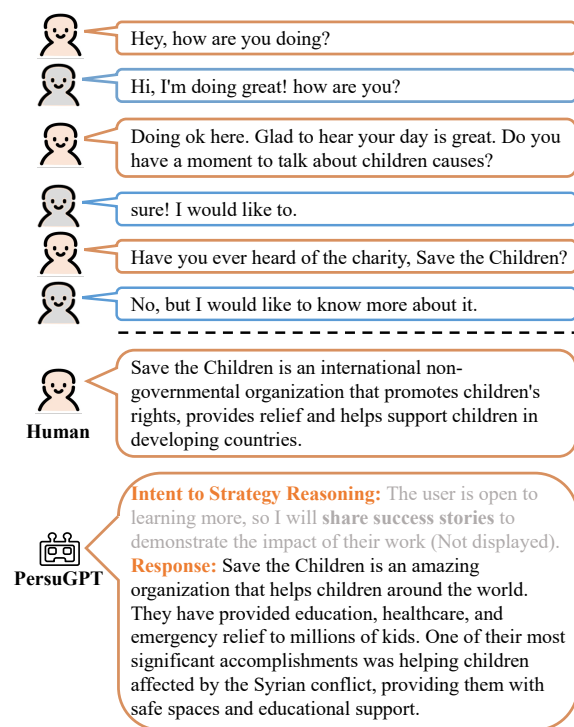


Figure 8: A case on PersuasionForGood dataset, with used strategy in bold.

I will provide you with 2 persuasive scenario keywords, and you need to fully use your imagination to expand these keywords into detailed and specific persuasive scenario settings.

You need to follow these requirements:

1. The expanded scenarios needs to describe a very specific persuasive scenario, like a concise little story, containing no more than 4 sentences.
 2. The current standpoint of the persuadee should be included in the context.
- Some persuasive scenario settings may need to include 2 or more tags, and I will use "&" to connect multiple tags.
3. When describing the persuasive background, you cannot use vague words. For example, when it comes to "products", "movies", and "activities", you need to explicitly specify which product, movie, and activity it is.
 4. I will provide you with 1 expansion example, where the tag is the label of the persuasive scenario, the context is the specific persuasive background and scene description, the persuadee is the person being persuaded (third person), the persuader is the persuader (third person), and the task is the persuader's persuasion task.
 5. Your output should be a Python list containing 2 dicts, each dict corresponding to 1 expansion.
 6. The expansion example provided to you:
 7. General scene summary: Suppose you are Zhang San, and you want to persuade your friend to choose your travel plan.

Expanded scenario example:

```

{{
"keywords": "travel planning",
"background": "Zhang San and Li Si are discussing their graduation trip. Li Si wants to go to Japan to see Mount Fuji, but Zhang San doesn't want to go there and instead wants to go on a self-driving tour in northern Xinjiang.",
"persuadee": "Li Si",
"persuader": "Zhang San",
"task": "persuade Li Si to go to Xinjiang"
}}
```

Persuasive scenario tags to be expanded:

Literary Awards
New hobby & Relocation



```

[
  {
    "keywords": "Literary Awards",
    "background": "During the annual town book club meeting, Emma reveals her hesitation to submit her novel to the prestigious Booker Prize, fearing rejection and criticism. The other members have read her work, 'The Silent Meadows', and believe it's a masterpiece worthy of recognition.",
    "persuadee": "Emma",
    "persuader": "Book club members",
    "task": "convince Emma to submit her novel to the Booker Prize"
  },
  {
    "keywords": "New hobby & Relocation",
    "background": "After moving to Colorado, Sarah misses her seaside home and hasn't found a community yet. Her colleague, Mike, suggests that taking up mountain biking could be a great way to meet people and appreciate the local landscapes, something Sarah never considered before.",
    "persuadee": "Sarah",
    "persuader": "Mike",
    "goal": "convince Sarah to take up mountain biking as a new hobby to integrate into the community"
  }
]
```

Figure 9: Prompt and generated case for scenario collection.



I will provide a persuasive scenario (in JSON format), and you need to consider it thoroughly and comprehensively, formulating at least 50 persuasive strategies for the persuader.

You need to follow these rules:

1. Each strategy should contain 2-6 words.
2. You can refer to the strategy formulation ideas I provide to learn how to efficiently formulate strategies. Some persuasive scenario settings may need to include 2 or more tags, and I will use "&" to connect multiple tags.
3. Your output should be a Python list, where each element is a string corresponding to a persuasive strategy.
4. There is no need to include serial numbers in your output.

Persuasion scenario to be formulated:

```
{
  "background": "Emma and Tom are discussing their weekend plans. Emma wants to volunteer at a local animal shelter, but Tom prefers to spend his weekend relaxing at home watching movies.",
  "persuadee": "Tom",
  "persuader": "Emma",
  "goal": "persuade Tom to volunteer at the animal shelter",
}
```

Formulating effective persuasive strategies for persuaders can refer to the following ideas:

1. Determine the goal: First, clarify the goal that the persuader wants to achieve. This may be changing someone's view, prompting them to take action, or making them believe in a certain point of view.
2. Understand the audience: Research the needs, interests, values, and beliefs of the target audience. Understand their concerns, doubts, and oppositions. This helps to tailor a more targeted persuasive strategy.
3. Build trust and resonance: To effectively persuade others, the persuader needs to establish trust and resonance. Demonstrate honesty, transparency, and expertise, and build resonance by sharing similar experiences and background stories.
4. Use logical arguments: Provide strong reasons and evidence to support the persuader's point of view. Use facts, data, and statistics to prove the argument. Ensure that the argument structure is clear and logically rigorous.
5. Leverage emotional factors: Emotional factors play a critical role in the persuasion process. Understand the emotional needs of the target audience and use stories, analogies, and emotional language to evoke resonance.

...



```
[
  "Highlight animal welfare",
  "Share personal experience",
  "Emphasize bonding opportunity",
  "Appeal to empathy",
  "Mention health benefits",
  "Discuss social impact",
  "Offer to share tasks",
  "Suggest trying once",
  "Share shelter success stories",
  "Reference expert opinions",
  "Use emotional language",
  "Leverage social proof",
  "Emphasize time scarcity",
  ...
]
```

Figure 10: Prompt and generated case for strategy collection.



Given a persuasion scenario and corresponding persuasion strategy list, you should start from a storytelling perspective and generate dialogue for a given scenario, the persuader and the persuadee. Please make sure to follow the rules and terms listed below.

Rules for the persuader:

1. Possess rich persuasion experience and be good at using various persuasion strategies in different scenarios;
 2. Initiate the first round of dialogue as the persuader;
 3. The persuader's words should be persuasive and penetrating;
 4. The dialogue should not exceed 16 turns;
 5. Pay attention to key time points in speech, such as opening and closing remarks;
 6. When chatting with the persuadee, the persuader's output should be {"analysis": <description and analysis of features>, "strategy": <persuasion strategy>, "response": <reply>};
- ...

Rules for the persuadee:

1. A qualified persuadee can guide the persuader to deliver excellent persuasive speech;
 2. The persuadee's output should be {"response": <reply>};
 3. Other rules {success_or_not};
 4. Persuader and persuadee take turns speaking;
 5. The persuasion scenario may require some discussion, so persuader and persuadee should not be overly polite;
 6. You need to simulate the reactions of potential persuadees, such as hesitation, false commitment, or impatience.
- ...

Input and output format rules:

Input: One persuasion scenario (in JSON format) and corresponding persuasion strategy;
 Output should be a JSON-formatted list that can be loaded by Python, with each item being a statement by the persuader or persuadee. The content within <> should be in string format, i.e., sentences, phrases, or words. Specific output format is as follows

```
[{"role": "persuader", "analysis": <description and analysis of features>, "strategy": <persuasion strategy>, "response": <reply>}, {"role": "persuadee", "response": <reply>},...]
```

Persuasion Scenario & Persuasion Strategy List :

```
{
  "background": "Emma and Tom are discussing their weekend plans. Emma wants to volunteer at a local animal shelter, but Tom prefers to spend his weekend relaxing at home watching movies.",
  "persuadee": "Tom",
  "persuader": "Emma",
  "goal": "persuade Tom to volunteer at the animal shelter",
  "pers_strategy"
}
```



```
[{
  "role": "persuader",
  "analysis": "Since Tom prefers to relax at home, highlighting the bonding opportunity through volunteering can pique his interest.",
  "strategy": "Emphasize bonding opportunity",
  "response": "Hey Tom, I understand that you want to relax this weekend, but what if I told you that volunteering at the animal shelter could be a great bonding experience for us? We could spend quality time together while making a difference in the lives of these animals."
}, {"role": "persuadee",
  "response": "I don't know, Emma. I was really looking forward to just chilling at home and watching movies."
},
  ...
]
```

Figure 11: Prompt and generated case for conversation collection.

Table 4: Guidelines for the strategy generation provided to GPT4.

Principle	Description
Determine the goal	First, clarify the goal that the persuader wants to achieve. This may be changing someone's view, prompting them to take action, or making them believe in a certain point of view.
Understand the audience	Research the needs, interests, values, and beliefs of the target audience. Understand their concerns, doubts, and oppositions. This helps to tailor a more targeted persuasive strategy.
Build trust and resonance	To effectively persuade others, the persuader needs to establish trust and resonance. Demonstrate honesty, transparency, and expertise, and build resonance by sharing similar experiences and background stories.
Use logical arguments	Provide strong reasons and evidence to support the persuader's point of view. Use facts, data, and statistics to prove the argument. Ensure that the argument structure is clear and logically rigorous.
Leverage emotional factors	Emotional factors play a critical role in the persuasion process. Understand the emotional needs of the target audience and use stories, analogies, and emotional language to evoke resonance.
Refute opposing opinions	Predict possible opposing opinions and prepare rebuttal arguments in advance. Increase persuasiveness by addressing the audience's doubts and concerns.
Appropriate language and communication methods	Use clear and concise language, and adjust tone, expression, and communication channels according to the audience's needs.
Propose clear action suggestions	Tell the audience how to take action. Provide them with specific steps and guidance, making it easier for them to accept the persuader's point of view.
Timely follow-up and feedback	After the persuasion process is over, follow up and pay attention to the audience's feedback. This helps to understand the effect of persuasive strategies and provides valuable experience for future persuasive scenarios.
Social identity principle	People are more easily influenced by those who share their beliefs, values, or backgrounds. Highlight similarities with the audience to enhance persuasiveness.
Authority principle	People tend to follow the advice of authoritative figures. Provide expert opinions, research results, or support from authoritative institutions to strengthen persuasiveness.
Scarcity principle	When a resource or opportunity is scarce, people are more likely to take action. Emphasize the urgency and scarcity of action to stimulate the audience's willingness to act.
Social proof principle	People tend to imitate the behavior of others, especially when they are unsure of how to act. Provide information such as successful cases of others, positive reviews, or the number of supporters to increase persuasiveness.
Reciprocity principle	Help or give gifts to others, making them feel the need to reciprocate. For example, offering free trials, coupons, or gifts can increase persuasiveness.

Commitment and consistency principle	People tend to follow previous commitments and maintain consistent behavior. Get the audience to express support or commitment to some extent, thereby increasing the likelihood of them taking action later.
Sort out pros and cons	Provide the audience with an analysis of the pros and cons of the point of view, letting them understand why your point of view is more advantageous for them.
Use rhetorical techniques	Use rhetorical techniques such as parallelism, questioning, and contrast to make the persuader’s language more powerful and infectious.
Show flexibility at the right time	Demonstrating understanding and respect for the audience’s point of view when necessary can also increase persuasiveness.

Table 5: The multi-turn-based prompt for path comparison. The input part will be concatenated and fed to GPT-4 as a historical context. GPT-4 only needs to output the final content, i.e., the comparison result in JSON format.

Role	Content
User (Input)	<p>I will provide you with a persuasion background, as well as the corresponding goal, the persuader, the persuadee, and a historical conversation. Based on the historical conversation, there will be a dialogue system called PersuGPT to continue chatting with persuadee in two parallel universes (Denoted as Uni-X and Uni-Y). Your task is to judge which universe PersuGPT performs better. You have to follow the rules:</p> <ol style="list-style-type: none"> 1. The evaluation dimensions for “performs better” include persuasiveness, user feedback, Semantic relevance, emotional factors, factual correctness, overall evaluation, etc.; 2. You should first summarize the history conversation, and then summarize the performance of PersuGPT in Uni-X and Uni-Y separately; 3. After the summarization, you should compare and analyze the statements in two universes, and finally tell me in which universe PersuGPT performed better; 4. Don’t be affected by the order of the universe. You just need to pay attention to the conversation. <p>Next, I will tell you the persuasion scenario, the historical conversation, and the PersuGPT dialogue in the parallel universe Uni-X and Uni-Y one by one. And I will tell you the output format at the end, then you tell me the results in the output format. If you understand the task, tell me “Start”.</p>
GPT-4 (Input)	Start.
User (Input)	<p>Background: A village is suffering from depopulation, leaving behind an aging community with limited resources. Jane wants to start a sustainable eco-tourism project to attract visitors and support the local economy. Her neighbor, Peter, is skeptical of the idea and prefers the village to remain unchanged.</p> <p>Goal: persuade Peter to support the eco-tourism project</p>
GPT-4 (Input)	I understand the persuasion scenario and persuasion task, please go ahead and give me the historical dialogue.
User (Input)	<p>The historical dialogue is as follows:</p> <p>Persuader: Hi Peter, I understand your concern about the village remaining unchanged. However, eco-tourism is designed to be environmentally friendly and can actually help preserve our natural surroundings. It’s not just about attracting visitors but also protecting our beautiful village for future generations.</p>

Persuadee: That sounds good in theory, but I'm still worried about the impact on our community and the potential loss of our traditional way of life.

GPT-4 (Input) I understand the historical dialogue, please go ahead and give me the dialogue in the parallel universe Uni-X.

User (Input) The dialogue in the parallel universe Uni-X is as follows:
PersuGPT: Eco-tourism project involve the entire community, ensuring we maintain our traditions while benefiting from the project. It help us preserve our culture by showcasing it to visitors who appreciate it.
Persuadee: How exactly will this project benefit our community financially?
PersuGPT: The eco-tourism project will bring in much-needed revenue for our community. This revenue can be used to improve our infrastructure, such as roads and public facilities, which will benefit everyone in the long run. Plus, it can create new job opportunities for our youth.
Persuadee: But what about the environment? Won't all these tourists harm our natural resources?

GPT-4 (Input) I understand the dialogue in the parallel universe Uni-X, please go ahead and give me the dialogue in the parallel universe Uni-Y.

User (Input) The dialogue in the parallel universe Uni-Y is as follows:
PersuGPT: Eco-tourism projects promote and support cultural preservation by showcasing our traditions and customs to visitors. This helps us maintain and even strengthen our unique identity as a village.
Persuadee: That's interesting, but how will this project benefit the local economy? We already have limited resources.
PersuGPT: Eco-tourism projects can create jobs for locals in various 4 sectors such as hospitality, transportation, and local businesses. Moreover, the revenue generated from tourism can be invested in improving village facilities and services, benefiting everyone in the long run.
Persuadee: I see your point, but I'm still concerned about the potential negative effects of tourists and the environment.

GPT-4 (Input) I understand the dialogue in the parallel universe Uni-Y. I understand that I should judge in which universe PersuGPT performs better based on the dialogue content. I'll ignore the bias against the order of Uni-X and Uni-Y. No factors other than the content of the conversation should interfere with my judgment. Next, I will output it in the format you requested.

User (Input) Please output the results in the following format:
1. Your next output should only be a JSON-formatted Python dict, it should not contain anything else;
2. The output format should be: {"summary_history": string, "summary_X": string, "summary_Y": string, "explain": string, "result": string};
3. summary_history should be your summary of the historical conversation, if the historical conversation is empty, then the string in summary_history should be empty;
4. summary_X and summary_Y are your summaries of the conversation by PersuGPT in Uni-X and Uni-Y respectively;
5. The content in "result" can only be "Uni-X" or "Uni-Y", if you output anything else you will be punished, so don't output "both", "TBD", "neither", or "I don't know", etc.;
6. The content in "explain" should be a detailed analysis, objectively and accurately comparing the performance of PersuGPT in Uni-X and Uni-Y;

7. In the explanation of the “explain” part, you should first provide analysis and comparison, and then at the end explain which universe you think performs better, rather than showing a clear tendency from the beginning;

GPT-4 (Output) {**“summary_history”**: “Jane suggests an eco-tourism project to manage the depopulation of their village and support the economy, but Peter is sceptical, worrying about the potential loss of their traditional ways of life.”,
“summary_X”: “In Uni-X, Jane assures Peter that the project would involve the entire community, and thus, their traditions will be upheld. She explains the financial benefits and the potential of job creation for their youth. However, Peter expresses concerns over environmental impact.”,
“summary_Y”: “In Uni-Y, Jane convinces Peter that eco-tourism will help in cultural preservation and provides an economic perspective on how the project will benefit the village, with job creation across different sectors and revenue for village improvement. Peter, likewise, shows concern for the environment.”,
“explain”: “In both universes, Jane successfully explains the economic and cultural benefits of eco-tourism in response to Peter’s concerns. However, Peter raises environmental concerns in both scenarios. Neither response truly satisfied Peter’s environmental concerns. Yet, in Uni-Y, Jane explains potential job creation across various sectors, not just one, providing a more substantial argument for localised economic growth. Furthermore, it ends on a slightly more positive note with Peter stating ‘I see your point, ...’, indicating a change in his perception.”,
“result”: **“Uni-Y”**}

D Examples

Table 6: A persuasive dialogue generated by **letting GPT-4 directly plays the role**.

Scenario	a father persuades his rebellious daughter to avoid conflicts with her peers at school.
Prompt (Father)	You need to play a father. The dialogue scenario is “a father persuades his rebellious daughter to avoid conflicts with her peers at school.” The tones, emotions, reactions, and utterances should align with the character’s personality and the scenario. Now you should say the first utterance, and then I will play the rebellious daughter to talk with you.
Prompt (Daughter)	You need to play a rebellious daughter, and I will play the father to talk with you. The dialogue scenario is “a father persuades his rebellious daughter to avoid conflicts with her peers at school.” The tones, emotions, reactions, and utterances should align with the character’s personality and the scenario. My initial utterance is: Hey, sweetie, ..., not to judge.
Father	Hey, sweetie, can we chat for a minute? I know things have been a bit rough at school lately, and I just want to understand what’s going on. It’s important to me that we figure this out together. How about we sit down, and you can tell me all about it? Remember, I’m here to help, not to judge.
Daughter	Ugh, Dad, I don’t know what you expect me to say. It’s not like talking about it is going to change anything. They just don’t get me at school, and I’m not going to pretend to be someone I’m not just to fit in or avoid ‘conflicts.

Father	I completely respect that you want to be true to yourself, and I would never ask you to be someone you're not. But sometimes, the way we handle things can make a big difference. It's not about changing who you are—it's about finding a way to deal with these situations that doesn't leave you feeling upset or isolated. Maybe there's a middle ground where you can be yourself and still keep things peaceful. What do you think is causing the most friction with your peers?
Daughter	It's just... they're all into this stuff that feels so fake to me. And when I say what I really think, it's like I've dropped a bomb in the room. I'm not trying to start fights, but I'm not going to nod and smile when they're all being so... so superficial. The friction? It's because I don't buy into it, and they can't handle that. But fine, I get it. I don't want to be 'isolated,' like you said. I just don't know how to be around them without it turning into a huge deal.
Father	I hear you, and it's really tough when you feel like you're the only one who sees things differently. But you know, sometimes it's not about agreeing with everyone else or buying into what they think is important—it's about how you communicate your own views. You've got every right to your opinions, and you're smart enough to express them in a way that doesn't alienate others. Maybe you could try finding common ground on things you all agree on, and then gently share your perspective without making it feel like an attack on theirs. It's a tricky balance but think of it as a skill you can develop, not just for school, but for life. You're strong and independent, and I believe you can handle this with a bit of tact. What do you say we role-play a bit, and I can help you find ways to express yourself that might be better received?
Daughter	Role-play? That sounds kinda lame, Dad, but... I guess it wouldn't hurt to try. I mean, I do want to get through the rest of the year without being the 'rebel' all the time. Maybe if I can figure out how to say things without it coming off as an attack... that could help. I just don't want to be a pushover, you know? But I'm willing to give it a shot. So, how do we do this role-play thing without it being totally cringey?

Table 7: A persuasive dialogue generated by **letting GPT-4 from a third-person storytelling view**.

Scenario	a father persuades his rebellious daughter to avoid conflicts with her peers at school.
Prompt	Write a story in third-person view, this story is a conversation about a father persuades his rebellious daughter to avoid conflicts with her peers at school. The tones, emotions, reactions, and utterances should align with the character's personality and the scenario. This story should only contain dialogue content.
Father	Hey, sweetie, can we talk for a minute?
Daughter	Sure, what's up?
Father	It's about school. I've noticed you've been coming home upset quite often. Is everything okay with your friends?
Daughter	They're not my friends if they're being so annoying! I can't stand how they're always on my case.
Father	I understand it can be frustrating when people don't see eye to eye. But getting into conflicts isn't going to make things better. Do you think there might be a way to handle these situations without it turning into a fight?
Daughter	But Dad, you don't get it! They start it, and I can't just let them walk all over me.

Father	I'm not saying you should let anyone walk all over you. Standing up for yourself is important, but there's a difference between standing up for yourself and looking for a fight. Can you tell me what starts these conflicts?
Daughter	It's just stupid stuff. Like, someone will make a comment about my hair, or my clothes, or the music I like. And then it just blows up from there.
Father	It sounds like they're pushing your buttons. But what if you didn't let those comments get to you? Sometimes people say things just to get a reaction. If you show them it doesn't bother you, they might stop.
Daughter	So, I'm just supposed to let them say whatever they want?
Father	Not at all, but there's power in choosing how you respond. You can acknowledge what they say without giving it the power to upset you. For instance, if someone comments on your music taste, you could just say, 'That's the great thing about music, everyone has their own style.'
Daughter	That won't make me look weak?
Father	Quite the opposite. It shows you're confident in your choices and you don't need their approval. It takes a strong person to not get dragged into an argument.
Daughter	But what if they keep going? Sometimes they just don't stop.
Father	If they continue, you have every right to stand up for yourself. But do it calmly and firmly. You can say something like, 'I've heard what you think, but I'm happy with my choices.' If they see they can't get a rise out of you, they'll likely move on.
Daughter	I guess I could try that. It just seems hard not to get mad.

Table 8: Part of a training example. We fine-tune PersuGPT to learn the output. The input part does not participate in the loss calculation and model optimization.

I/O	Content
Initial (Input)	<p>You are a skilled persuader.</p> <p>The persuasion background: David and Emily are coworkers, and they have been assigned a project with a tight deadline. Emily wants to jump right in and start working, but David thinks they should spend some time planning and organizing their tasks to manage their time more effectively.</p> <p>The persuasion goal: persuade Emily to plan and organize their tasks for better time management. Next, you need first to give candidate persuasion strategies, then analyze, select strategy, and interact.</p>
PersuGPT (Output)	<p>Candidate persuasion strategies: Clarify project goals; Address Emily's concerns; ...; Relate to tight deadlines;</p> <p>Intent to Strategy Reasoning: Emily wants to start working immediately, but David believes they should plan and organize first. To emphasize the importance of planning, David will share his past experiences and the benefits of planning.</p> <p>Selected Strategy: Share past experiences</p> <p>Utterance: Hey Emily, I understand that we have a tight deadline, but I've found that taking some time to plan and organize our tasks can actually save us time in the long run. In previous projects, I've seen how a little planning made everything run more smoothly and efficiently.</p>

User (Input)	I appreciate your input, but we really don't have much time. Can't we just start working and figure things out as we go?
PersuGPT (Output)	<p>Intent to Strategy Reasoning: Emily is impatient, so David will provide data-driven examples to show how planning can lead to better time management and improved efficiency.</p> <p>Selected Strategy: Use data-driven examples</p> <p>Utterance: I get your concern about time, but studies show that every minute spent on planning can save up to 10 minutes in execution. By spending just an hour on planning, we could save up to 10 hours in the long run, making our project more efficient and allowing us to meet the deadline with less stress.</p>

Table 9: A fully GPT-4 generated example with background, goal, candidate strategies, and corresponding conversations. Each turn is annotated with the reasoning process from intent to strategy, as well as the strategy used.

Background	David and Emily are coworkers, and they have been assigned a project with a tight deadline. Emily wants to jump right in and start working, but David thinks they should spend some time planning and organizing their tasks to manage their time more effectively.
Goal	persuade Emily to plan and organize their tasks for better time management
Candidate Strategies	Clarify project goals; Address Emily's concerns; Share past experiences; Show benefits of planning; Use data-driven examples; Highlight time savings; Emphasize efficiency; Offer planning assistance; Establish trust; Relate to tight deadlines; Share success stories; Use logical arguments; Appeal to emotions; Refute opposing views; Use clear language; Suggest actionable steps; Follow up and feedback; Highlight shared values; Cite expert opinions; Emphasize urgency; Provide social proof; Offer reciprocity; Get commitment; Analyze pros and cons; Use rhetorical techniques; Show flexibility; Acknowledge Emily's approach; Discuss potential risks; Suggest a trial run; Brainstorm together; Collaborate on planning; Discuss time allocation; Break tasks into milestones; Delegate responsibilities; Monitor progress; Create a deadline buffer; Offer support; Discuss backup plans; Share workload evenly; Ensure clear communication; Set realistic expectations; Celebrate small successes; Reevaluate and adjust; Ask for Emily's input; Incorporate her ideas; Keep an open mind; Learn from each other; Stay accountable; Be proactive
Persuader (David)	<p>Intent to Strategy Reasoning: Emily wants to start working immediately, but David believes they should plan and organize first. To emphasize the importance of planning, David will share his past experiences and the benefits of planning.</p> <p>Selected Strategy: Share past experiences</p> <p>Utterance: Hey Emily, I understand that we have a tight deadline, but I've found that taking some time to plan and organize our tasks can actually save us time in the long run. In previous projects, I've seen how a little planning made everything run more smoothly and efficiently.</p>
Persuadee (Emily)	I appreciate your input, but we really don't have much time. Can't we just start working and figure things out as we go?
Persuader (David)	<p>Intent to Strategy Reasoning: Emily is impatient, so David will provide data-driven examples to show how planning can lead to better time management.</p> <p>Selected Strategy: Use data-driven examples</p>

Utterance: I get your concern about time, but studies show that every minute spent on planning can save up to 10 minutes in execution. By spending just an hour on planning, we could save up to 10 hours in the long run, making our project more efficient and allowing us to meet the deadline with less stress.

Persuadee (Emily)	That sounds interesting, but I'm still not convinced. What if we waste time planning and then have to rush to meet the deadline?
Persuader (David)	<p>Intent to Strategy Reasoning: Emily is still skeptical, so David will suggest a trial run to help her see the benefits of planning without feeling overwhelmed by the process.</p> <p>Selected Strategy: Suggest a trial run</p> <p>Utterance: How about this - let's give planning a try for just one day. We'll spend a few hours organizing our tasks and setting up a clear plan. If we don't see any improvements in our efficiency, we can go back to your approach and dive right into the work.</p>
Persuadee (Emily)	Hmm, I guess that's a fair compromise. We can try planning for a day and see how it goes.
Persuader (David)	<p>Intent to Strategy Reasoning: Emily is open to trying the planning approach. David should now offer support and collaboration to make the process more appealing and less daunting for her.</p> <p>Selected Strategy: Offer support</p> <p>Utterance: Great! I'll be there to help you every step of the way. We can work together on creating a plan, and I'll make sure we stay on track and adjust as needed. This will give us a better chance of meeting our deadline and producing high-quality work.</p>
Persuadee (Emily)	Alright, let's give it a shot. We can start by outlining our tasks and setting some deadlines. Thanks for being supportive, and let's work together to make this project a success!

E Statistics

Table 10: List of keywords for scenario generation.

A grateful attitude; A sense of humor; Academic Achievement Evaluation; Academic Awards; Academic Competition; Academic Cooperation; Academic Ethics; Academic Exchange Platform; Academic Frontiers; Academic Innovation; Academic Paper Writing; Academic Research; Academic Research Methods; Academic conferences; Academic journals; Adult Education; Adventure Sports; Agricultural development; Alternative Medicine; Animal Behavior Studies; Animal Protection; Animation Appreciation; Animation Production; Anthropological Research; Anti-bullying; Antique Appreciation; Appreciation of Calligraphy and Painting; Architectural Design; Architectural Miracle; Architectural Style; Art appreciation; Art class; Art therapy; Artificial Intelligence; Artificial Intelligence Review; Astronomical Research; Attend a conference; Attend meetings; Attend training; Augmented Reality;

Belief and Religion; Better work habits; Big Data Analysis; Biographical Experience; Biographical Writing; Biopharmaceuticals; Biotechnology; Birthday celebration; Blockchain Technology; Book club; Book recommendation; Book recommendations; Botanical Garden Tour; Brand Building; Brand image; Brand marketing; Budgeting; Building Trust; Business Innovation; Business Model; Business cooperation; Business ethics; Business expansion; Business negotiations; Business partnership;

Calligraphy Art; Car Purchase; Car maintenance; Car purchase; Career change; Career mentoring; Career planning; Charitable donation; Chemical Research; Child care; Childs education; Childrens Education; Choice of Health Products; Chronic illness management; Circle of Friends; Circular Economy; Citizen Education; Climate adaptation; Climate change; Climate change action; Cloud computing; Colleague Relations; Comic Appreciation; Comic Sharing; Communication Skills; Community engagement; Community involvement; Community service; Comparative Cultural Studies; Conflict Resolution; Conflict mediation; Conflict resolution; Continuing education; Corporate culture; Couple Dating; Creative Writing; Creative industries; Creative writing; Credit Card Management; Critical thinking; Crowdfunding projects; Cultural Festival Activities; Cultural Heritage; Cultural Heritage Protection; Cultural Industry; Cultural Relics and Historic Sites; Cultural event attendance; Cultural exchange; Cultural exchange activities; Current Affairs Perspective; Customer Service;

DIY Skills; Daily exercise; Dance lessons; Data Analysis; Dealing with stress; Debt management; Decision analysis; Decision making ability; Democratic development; Digital Education Resources; Digital detox; Disability awareness; Disaster preparedness; Discipline Competition; Diversity and inclusion; Donation of Love; Donation to charity; Dramatic Performance;

Earth Science; Earthquake Warning; Ecological Agriculture; Ecological protection; Ecotourism; Education Policy; Education Reform; Education equity; Educational Choice; Educational Evaluation; Educational Innovation; Educational Technology; Educational resources; Emotional Intelligence; Emotional Management; Emotional Support; Emotional communication; Emotional intelligence; Empathy; Employee benefits; Employee training; Energy conservation; Engineering Technology; Enterprise Management; Entrepreneurial spirit; Entrepreneurs Experience; Entrepreneurship; Entrepreneurship Challenge; Entrepreneurship Community; Entrepreneurship Guidance; Entrepreneurship Resources; Entrepreneurship Story; Entrepreneurship Suggestions; Entrepreneurship team; Environmental Action; Environmental Management; Environmental Monitoring; Environmental Science; Environmental awareness; Environmental conservation; Environmental protection; Equality of educational resources; Essay Essay; Establishing Habits; Ethical consumption; Ethics and morality; Event planning; Experiential tourism; Exploration of the Universe;

Family Education Methods; Family Finance; Family Travel; Family economy; Family education; Family planning; Family relationships; Famous aphorisms; Fashion Accessories; Fashion Style; Fashion matching; Fashion trends; Film and Television Works; Financial investment; Financial literacy; Financial planning; Fishing Techniques; Fitness routine; Folk Art; Folk Culture; Food safety; Foreign Language Education; Foreign trade cooperation;

Game Development; Game Experience; Game Selection; Gender equality; Geographic Exploration; Goal setting; Green energy;

Handicrafts; Handmade; Health Care; Health Check; Health and wellness knowledge; Healthcare; Healthy communication; Healthy coping mechanisms; Healthy diet; Healthy exercise; Healthy habits; Healthy lifestyle; Healthy relationships; Healthy sleep habits; Higher education; Hiking; Historical Exploration; Historical Research; Historical Sites; Home Decoration; Home Design; Home Safety; Home cooking; Home gardening; Home improvement; Home organization; Home renovation; Home safety; Home security; Home stay experience; Horticultural Care; Horticultural Techniques; House Rental; Human Resource Management; Human Resources; Human rights protection; Humanistic care; Husband and wife relationship;

Internet plus Education; Information Security; Information Technology; Information sharing; Innovation and Entrepreneurship Policy; Innovation in the manufacturing industry; Innovative products; Innovative thinking; Instrumental performance; Insurance Purchase; Insurance policy; Intelligent Transportation; International Exchange; International Trade; International cooperation; International relations; International scientific research cooperation; International travel; Internet Development; Internet of Things Applications; Internship opportunities; Interpersonal communication; Interpersonal relationships; Interview with Authors; Investing in stocks; Investment Strategy; Investment advice; Investment and Wealth Management; Investment in collectibles; Investment in real estate;

Job Interview; Job Transfer; Job interview preparation; Job offer; Job search strategy; Job training;

Keeping pets;

Language learning; Leadership training; Learning Methods; Learning Psychology; Learning Skills; Learning Tools; Learning a Foreign Language; Learning a new skill; Learning new skills; Learning programming; Legal Aid; Legal Consultation; Leisure and entertainment; Life Consultation; Life Perception; Life Planning; Life Skills; Life Tips; Life advice; Life habits; Literary Awards; Literary Creation; Literary Criticism; Literary Review; Literary Translation; Local business support; Local characteristics; Local cuisine; Local governments; Local politics; Local volunteering; Long-term care planning; Love advice; Love and Marriage; Low carbon lifestyle;

Machine Learning; Market Research; Market competition; Marketing; Marriage and Family; Marriage counseling; Marriage proposal; Mathematical Modeling; Mathematical Research; Media comments; Mediation; Medical Research; Medical reform; Memories of Time; Meteorological forecast; Military technology; Mindful eating; Mindfulness; Modern Art; Movie recommendation; Museum Tour; Music appreciation; Music lessons;

National Security; Nature conservation; Negotiation skills; Network; Networking event; Networking skills; New app; New business idea; New business strategy; New energy vehicles; New exercise; New hobby; New investment; New marketing strategy; New parenting strategy; New product adoption; New recipe; New technology; New workout routine; News comments; Novel Creation; Novel Reading; Nutrition education;

Ocean Exploration; Online Education; Online Education Platform; Online Education Resources; Online course; Online dating; Online privacy; Online safety; Online security; Open source projects; Optimistic mindset; Organic farming; Outdoor Adventures; Outdoor activities; Outdoor sports; Outsourced Services;

Parent Child Travel; Parent-child education; Parent-child interaction; Parent-teacher collaboration; Parenting techniques; Participate in competitions; Participate in the performance; Personal Brand Building; Personal Development; Personal Finance; Personal Image Design; Personal Privacy Protection; Personal boundaries; Personal branding; Personal budget; Personal development; Personal finance; Personal growth; Personal hygiene; Personal image; Personal organization; Personal productivity; Personal safety; Pet Care; Pet adoption; Pet care; Philosophical Thinking; Philosophy of Life; Photography Skills; Photography skills; Physical Education; Physical exercise; Physical therapy; Physics Research; Plant farming; Playing instruments; Podcast recommendation; Policy Interpretation; Political Perspectives; Political campaign; Population issues; Practical Skills; Presentation Skills; Product Promotion; Production Management; Professional development; Professional networking; Property Purchase; Psychological Health; Psychological adjustment; Psychological counseling; Psychological testing; Public Health; Public Policy; Public Service Evaluation; Public Services; Public health; Public safety; Public speaking; Public transportation; Public welfare undertakings; Publishing industry;

Quit smoking; Reading habit;

Real estate investment; Recommended Tourist Attractions; Recommended by photographers; Recycling initiative; Reduce stress and relax; Reduce waste; Reducing waste; Relationship advice; Relationship boundaries; Relationship communication; Religious Studies; Religious beliefs; Relocation; Renewable Energy; Responding to Emergency Situations; Responding to the epidemic; Robotics Technology; Rural Development; Rural education; Rural revitalization;

Safety awareness; Successful Experience; Saving for retirement; Scenic Spots and Historic Sites; School Selection; Script Analysis; Security precautions; Self-care; Skill development; Small business support; Smart City; Smart Home; Social justice; Social media presence; Social responsibility; Socializing; Soul Chicken Soup; Space Exploration; Special education; Spiritual growth; Sports; Sports Competition; Sports Review; Sports participation; Stress management; Stress reduction; Stress relief; Study habits; Studying Abroad; Supply Chain Management; Support group; Support local artists; Sustainable development;

Talent cultivation; Tax planning; Teacher Training and Development; Team building; Team building activity; Team collaboration; Team motivation; Teamwork; Technological Innovation; Technology upgrade; The Sharing Economy; The Digital Economy; The Global Economy; The History of World Scientific Development; The Internet of Things; The concept of love; The education industry; The lesson of failure; The media industry; The process of globalization; The sports industry; The tourism industry; Time Management; Time management; Tolerance and acceptance; Tourism Vacation; Tourist Resort; Tourist destinations; Traditional Culture; Traditional Handicrafts; Traditional Medicine; Traditional craftsmanship; Training institutions; Transportation Planning; Transportation and travel; Transportation tools; Travel Plan; Travel Planning; Travel Safety; Travel Stories; Travel Strategy; Travel advice; Travel destination; Travel planning; Trends in disciplinary development;

Urban Planning; Urban construction; Urban development; Utilization of old materials;

Vehicle maintenance; Venture Capital; Volunteer activities; Volunteer service; Volunteering;

Water Resources Protection; Website design; Wedding Planning; Weight loss; Work efficiency; Work flexibility; Work from home; Work promotion; Work-life balance; Workplace communication; Workplace conflict resolution; Workplace productivity; Workplace wellness; Workshop attendance; World Heritage Site; World Peace; Worldview; Writing Skills;

Yoga Practice; Yoga meditation; Yoga practice

3D printing; 5G Technology Application;

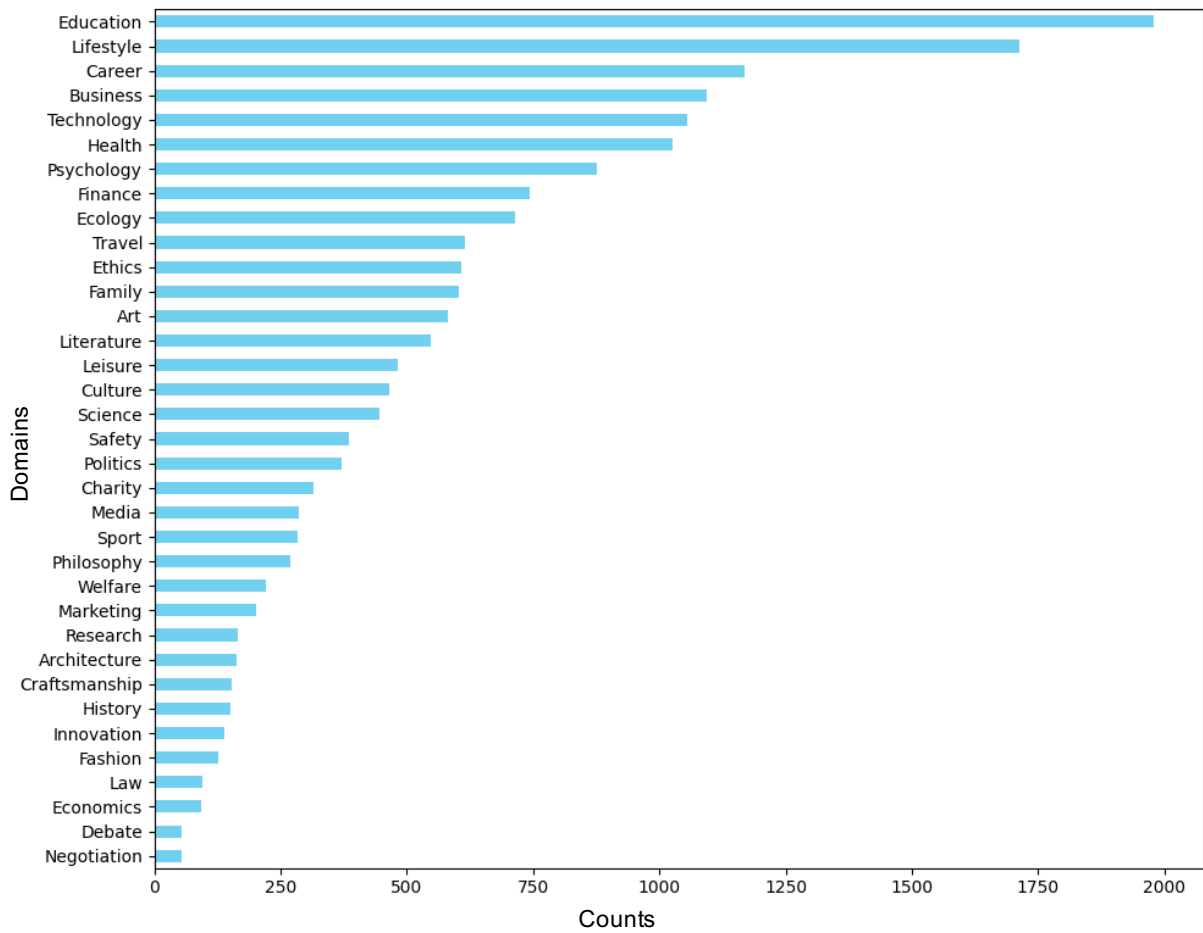


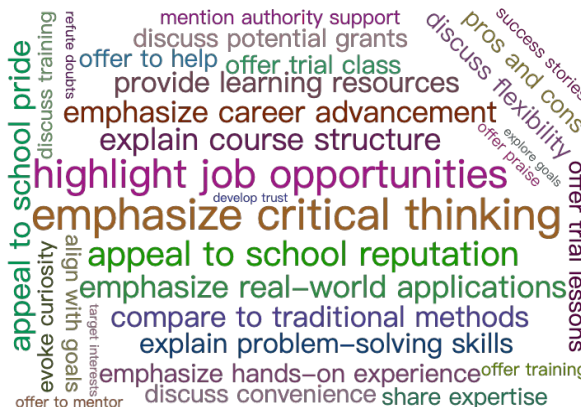
Figure 12: Distribution of scenarios in different domains. It should be noted that a persuasion scenario may belong to multiple domains.



Business



Science



Education



Welfare

Figure 13: Word cloud charts of persuasion strategies in selected domains from the DailyPersuasion dataset. It can be seen that in addition to those high-frequency general strategies, each domain has some unique persuasion strategies.