

PANDA: Empowering Small Language Models for Proactive Dialogue Through Agent-Based Synthesis (Student Abstract)

Rongyu Zhang¹, Dingyuan Zhang², Haopeng Li¹

¹Zhejiang University

²Hangzhou Dianzi University

rongyu.zhang@zju.edu.cn, ziink@hdu.edu.cn, haopeng.li@zju.edu.cn

Abstract

Proactive dialogue systems, which are designed to guide conversations toward predetermined goals. However, contemporary LLMs predominantly function as passive assistants, mechanically executing human instructions. A key challenge contributing to this limitation is the inherent difficulty in acquiring and annotating high-quality training data for proactive dialogue. Consequently, the scarcity of such data results in a notable deficiency in the proactive conversational capabilities of current LLMs. In this paper, we introduce **PANDA** (Proactive Agent-based Negotiation Dialogue Augmentation), a method designed to generate accurate, complex, and diverse proactive dialogue data for a challenging task—financial dispute mediation—where a LLM acts as the mediator. **PANDA** leverages a novel self-evolving synthesis process to manage a pool of user profiles and generate dialogues through structured interactions between multiple LLM-driven agents. To ensure data fidelity, we propose a comprehensive evaluation framework and build a two-level validation system combining automated and expert human verification. Our experiments demonstrate that an 8B-parameter model, trained on our synthesized dataset, achieves state-of-the-art results in the task’s evaluation framework. Its performance rivals top closed-source models guided by heavily engineered prompts, even when provided with only essential information.

Introduction

Proactive dialogue systems are vital in fields like finance, legal, and healthcare. Unlike traditional Q&A systems, these require strict adherence to processes and active information gathering. A key challenge is the lack of training data, as real-world data collection is costly and risky. Financial dispute mediation exemplifies the complexity of proactive dialogue, requiring models to follow legal procedures, manage emotions, and guide parties towards resolution.

To address this, we introduce PANDA, a framework for generating and validating financial mediation dialogue data via multi-agent collaboration. PANDA uses a self-evolving character pool and leverages both real data and large language model capabilities for diverse and authentic data synthesis. Through role play, knowledge graph retrieval, and su-

pervision, our system ensures procedural compliance and interaction authenticity.

Method

Figure 1 illustrates the comprehensive pipeline of our PANDA data synthesis method, which comprises three main stages. The efficacy of our method has been validated on models fine-tuned with this data.

Persona Composition

To ensure data diversity and authenticity, we initialize a persona pool from raw mediation dialogue corpora. After data cleansing and de-identification, we employ an LLM to automatically extract and refine core disputant attributes, such as personality (stubbornness, cooperation), professional background, family situation, and core claims. This collection forms our seed pool, which dynamically expands during generation via a self-evolving mechanism, providing rich and realistic “character scripts” for next-stage synthesis.

Multi-agent Interactive Dialogue Generation

Dialogue generation is accomplished through structured interaction among three LLM-driven agents, maximizing data authenticity and diversity.

Disputant Agent is assigned a persona from the pool, which dictates its linguistic style, emotional expression, and negotiation strategy. For instance, a “stubborn and distrustful” agent frequently challenges the mediator, while an “anxious but cooperative” one is more inclined to articulate challenges and seek solutions. This role-based approach ensures diverse interactive dynamics.

Mediator Agent, central to our framework, utilizes Retrieval-Augmented Generation to integrate with a financial mediation knowledge graph (Edge et al. 2024). This knowledge graph contains legal statutes, standard scripts, and compliant solutions. Accessing this repository ensures the agent provides professionally accurate information when explaining legal risks or proposing solutions, guaranteeing content reliability.

Evaluator Agent serves a supervisory role, verifying the execution of critical steps by the Mediator Agent and issuing subsequent instructions. It enforces dialogue quality by ensuring the Mediator maintains an anthropomorphic tone

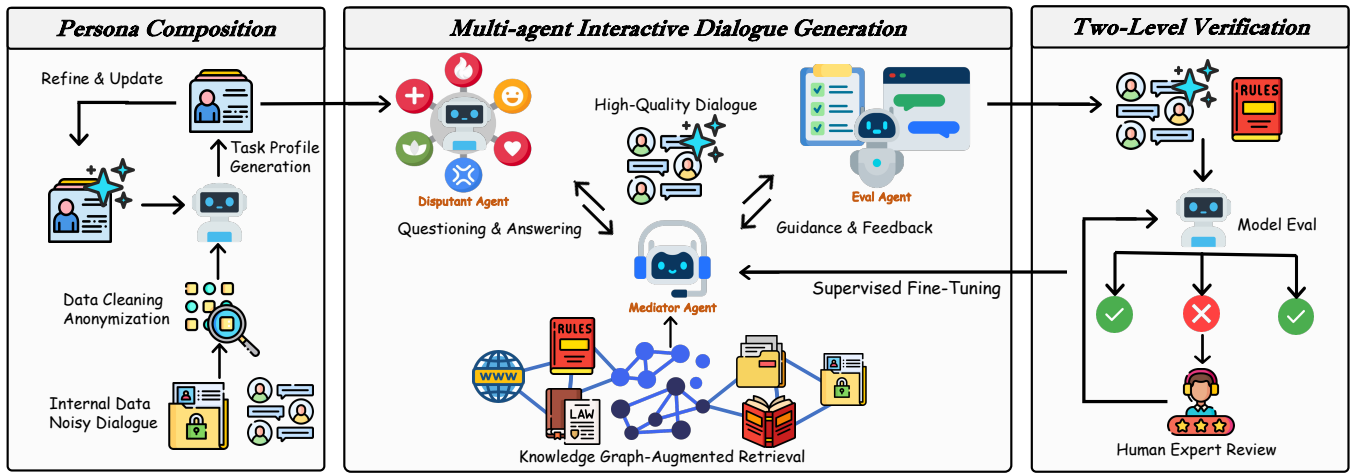


Figure 1: PANDA’s entire process consists of three main stages.

and conversational focus, thus mitigating topical drift and artificial responses.

Two-Level Verification

We implement a rigorous two-level verification process to ensure data quality. First, in collaboration with professional financial mediators, we define a seven-dimensional evaluation framework: Process Normativity, Script Precision, Emotional Suitability, Linguistic Clarity, Mediation Efficiency, Outcome Achievement, and Neutrality and Fairness. An LLM-based evaluator performs an initial screening to filter out low-quality dialogues. Subsequently, human experts review a subset of the data, providing feedback to refine the generation process. In a sample evaluation of 200 dialogues, human experts yielded an acceptance rate of over 98%, confirming the high quality and fidelity of data generated by the PANDA framework.

Experiments and Analysis

We evaluated the models using the “LLM as a Judge” methodology (Zheng et al. 2023), with Gemini 2.5 Pro serving as the impartial assessor. Our prompt-based evaluation framework covers seven expert-verified dimensions, each supported by few-shot examples and assigned specific weights to total 100 points.

In the evaluation, each model acted as a mediator conversing with an autonomous disputant agent. To simulate a realistic environment, no external knowledge bases were provided. Baseline models received detailed system prompts outlining mediation processes and strategies, whereas our model was only given the disputant’s information. All final scores are an average of 50 dialogue evaluations.

As shown in Table 1, our model (PANDA) demonstrates a significant advantage in procedural normativity, adhering more closely to standard mediation steps. While its performance in emotional suitability and language clarity is surpassed by GPT-4o, it outperforms other models of a similar parameter scale. Notably, PANDA achieves the highest overall score, exceeding even the leading closed-source model.

Evaluation Metric	Qwen3-8B w/ prompt	GPT-4o w/ prompt	PANDA w/o prompt
Procedural Normativity	10.22	14.86	21.16
Speech Precision	7.31	9.90	10.38
Emotional Suitability	5.29	9.71	7.80
Language Clarity	4.76	8.67	7.52
Mediation Efficiency	17.41	18.98	19.58
Outcome Achievement	6.57	9.04	10.64
Neutrality & Fairness	7.55	10.00	8.60
Overall Score	59.11	80.62	85.69

Table 1: Detailed performance comparison across all evaluation metrics.

Conclusion

This paper introduces PANDA, an innovative multi-agent-based data synthesis framework designed to address the scarcity of proactive dialogue data by generating high-quality conversational datasets. This work demonstrates that advanced data synthesis strategies can effectively empower smaller models to undertake complex and specialized proactive dialogue tasks traditionally reserved for larger models, paving the way for broader deployment and application of this technology in real-world scenarios.

References

- Edge, D.; Trinh, H.; Cheng, N.; Bradley, J.; Chao, A.; Mody, A.; Truitt, S.; Metropolitan, D.; Ness, R. O.; and Larson, J. 2024. From local to global: A graph rag approach to query-focused summarization. *arXiv preprint arXiv:2404.16130*.
- Zheng, L.; Chiang, W.-L.; Sheng, Y.; Zhuang, S.; Wu, Z.; Zhuang, Y.; Lin, Z.; Li, Z.; Li, D.; Xing, E.; et al. 2023. Judging llm-as-a-judge with mt-bench and chatbot arena. *Advances in neural information processing systems*, 36: 46595–46623.