Enhancing LLMs in Predictive Political QA with Semi-Structured Data

Yinan Liu[*,†] **Zihan Zhou**[†]
Zichun Jin **Xinyu Wang** **Bin Wang** **Xiaochun Yang**
School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China
[†]Equal contribution. [*]Corresponding author.

Abstract

Predictive political question answering (QA), such as predicting how a political actor will vote, goes beyond factual lookup. External political resources offer rich historical evidence, but rarely contain the answer itself. Existing LLM augmentation methods, including actor-profile-based simulation and knowledge graph evidence injection, improve political reasoning but largely treat external resources as knowledge-based evidence, leaving prediction-relevant signals under-modeled. We identify two complementary signals for predictive political QA: actor stances that capture issue-specific preferences, and high-order structure signals that capture indirect dependencies among political actors. We propose PSL, a dual-view framework that converts semi-structured political records into inference-oriented evidence for LLMs. PSL extracts stance signals from question-relevant actor records in a semantic view, and learns structure-aware actor representations from an actor interaction graph in a vector view. Across three real-world datasets and multiple LLMs, PSL consistently outperforms baselines, with ablations confirming the complementary gains of stance and structure signals.

1 Introduction

Political question answering (QA) serves diverse information needs about political actors, policies, and events. Large language models (LLMs) provide a powerful foundation for political QA, but their knowledge is often incomplete and unreliable in specialized, long-tail settings (Sahoo et al., 2024). Augmenting LLMs with external political resources is therefore a natural direction (Cuconasu et al., 2024). The political domain contains abundant external data, including social background information[1], electoral records[2], and legislative voting records[3], much of which is semi-structured or structured. However, many political QA tasks are not simple factual queries but predictive questions, such as predicting how a political actor will vote or respond to a future event (Yang et al., 2020; Galla and Burke, 2018). For these questions, external resources provide valuable records but rarely contain the answer directly, because the relevant behavior has not yet occurred and existing records serve only as indirect facts. Therefore, predictive political QA cannot be solved by direct factual retrieval alone.

To the best of our knowledge, the closest recent studies are PAA (Li et al., 2025) and PEG (Mou et al., 2024), which enhance LLMs' political reasoning with external resources. PAA uses actor-profile-based simulation, organizing actors' backgrounds and past behavior into textual profiles and prompting LLMs to simulate their decisions through role-playing. PEG uses knowledge-graph-based evidence injection, converting external resources into a knowledge graph (KG) and retrieving relevant structured evidence to augment LLMs. Both show the value of external resources for political reasoning, but still largely treat them as knowledge-based evidence (*e.g.* Political Behavior Profiles in Fig. 1). Moreover, profile-based methods struggle to capture implicit group-level influence among political actors (Sowden et al., 2018), whereas KG-based methods often lose fine-grained context when extracting raw resources into triples (Wang and Han, 2025). What remains missing is an evidence representation suited to predictive political reasoning, beyond treating external resources merely as knowledge-based evidence.

In this work, we identify two complementary signals encoded in political data. i) Actor stance signals, *i.e.*, their subjective attitudes toward specific issues, which reflect their values, ideology, or partisan interests (Burnham, 2025). Although implicit

[1]https://www.wikipedia.org/
[2]https://ballotpedia.org/
[3]https://legiscan.com/

in complex contexts, these signals possess greater reasoning potential compared to factual knowledge. ii) High-order structure signals, *i.e.*, the indirect dependencies within political actor interaction networks, which aim to characterize group influence. Political actors are connected not only by direct ties but also through higher-order relational patterns. For example, an actor's stance may influence others through intermediary organizations or decision-makers, forming an indirect network of policy influence. Such network-embedded dependencies are difficult to express in natural language, often leading to verbose descriptions and noise. Semi-structured data formats provide a suitable basis for modeling these two signals, as they preserve contextual detail while retaining the interactions among political actors encoded in the structure.

In this work, we propose PSL, a dual-view framework that uses semi-structured political data to enhance LLMs for predictive political QA. PSL extracts and integrates two complementary types of signals: i) actor stance signals from a semantic view and ii) high-order structure signals from a vector view. Specifically, PSL constructs an actor profile in JSON format for each political actor from abundant records, and links actors via shared records to build an interaction graph. In the semantic view, PSL retrieves question-relevant records from profiles and infers extensible stance signals to capture an actor's preference on a specific issue. In the vector view, PSL propagates information over the interaction graph to obtain actor representations that encode high-order neighborhood signals. PSL then co-embeds the actor representation with the question-focus embedding, producing mutually refined vectors, and equips the LLM with structure-aware reasoning through lightweight fine-tuning. Finally, PSL integrates stances from the semantic view with collaborative representations from the vector view and injects them into the LLM, providing stronger inferential evidence for predictive political QA.

This work makes three contributions. First, we identify two complementary signals in semi-structured political records: actor stances and high-order structure signals. The former capture actors' preferences on specific issues, whereas the latter characterize group influence revealed by shared political behavior. Second, we propose PSL, a dual-view framework for these two signals: a semantic view extracts actor-stance signals, and a vector view models high-order structure signals, jointly enhancing LLMs' predictive reasoning. Third, we conduct comprehensive experiments on three real-world datasets across different LLMs. The results demonstrate that PSL significantly outperforms all the baseline methods. Ablations verify the complementary benefits of semantic stance evidence and vectorized structure signals.

2 The PSL Framework

As illustrated in Figure 1, PSL derives inference-friendly evidence from semi-structured data through a dual-view strategy. The semi-structured format provides foundational support for this design. Compared with KG triples, it preserves finer-grained context for stance extraction, while its structured fields support more precise retrieval and naturally enable the construction of an actor interaction graph. On this basis, PSL models political evidence from two views. Actor stances are expressed as natural-language judgments in the semantic view, whereas high-order structure signals are expressed as actor embeddings in the vector view. To enable joint use of these complementary signals, PSL introduces two levels of synergy. Within the vector view, inspired by collaborative filtering (Schafer et al., 2007), PSL co-embeds the actor and the question representations to produce mutually refined vectors, allowing the LLM to exploit structure signals in downstream reasoning. Across the semantic and vector views, it injects stance and structure signals through a hybrid prompt.

2.1 Actor Stance Acquisition

Political Behavior Profiles. We construct human-centered, domain-specific, and semi-structured political behavior profiles that encapsulate factual knowledge related to U.S. politics, aiming to facilitate political understanding without excessively compressing the underlying information. We first collect politically relevant data sourced from legislative and diplomatic records. Specifically, we obtain legislative data via the LegiScan API, which contains legislators' voting records and bill descriptions. For diplomatic events, we extract interaction records involving political actors, particularly those pertinent to U.S. politics (Boschee et al., 2015). We reorganize these raw data into a human-centered structure (*i.e.*, $file(P)$). Based on these data, we extract a set of political actors, denoted by P. For each actor $p \in P$, we construct a corresponding political behavior profile, denoted by $file(p)$ in

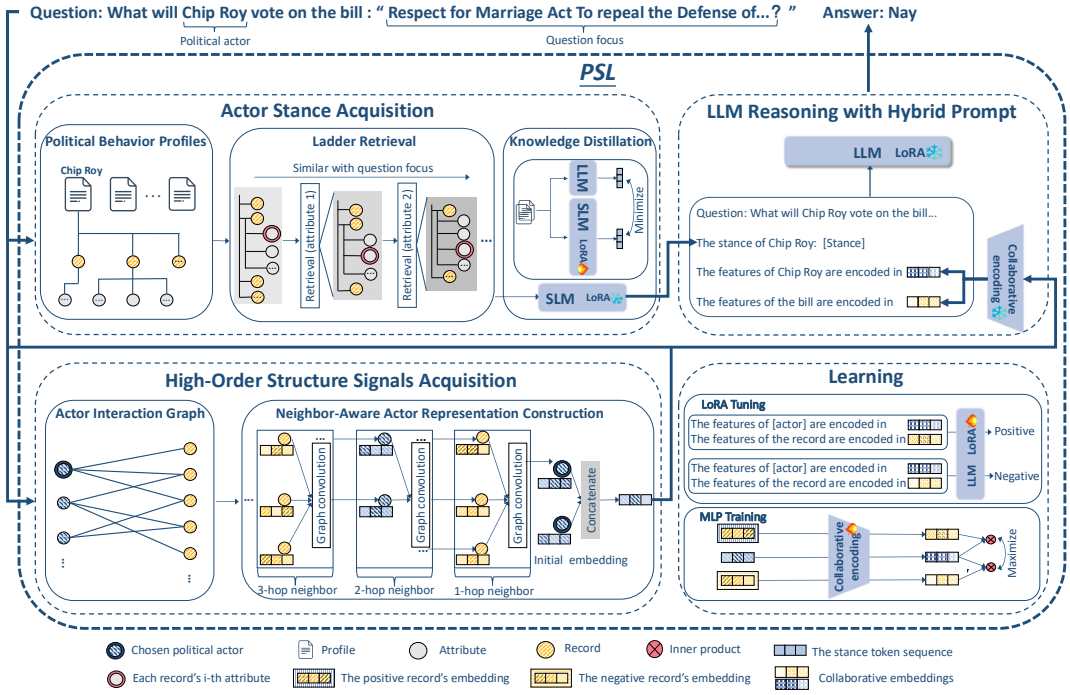


Figure 1: Overview of our framework PSL.

JSON format. The set of $file(p)$ is denoted as $file(P)$. Each profile $file(p)$ consists of a set of records, where each record $r_p \in file(p)$ is an object containing multiple attributes, and each attribute is expressed as $a_{r_p} \in r_p$, which is a key-value pair describing the actor's political behavior. For legislative events, each record r_p corresponds to a specific voting instance (*i.e.*, voting record), with attributes such as the bill's title, description, and the actor's vote. For diplomatic events, r_p denotes a concrete interaction event, with attributes detailing the action type and involved parties. We summarize bill descriptions with a summarization model (Lewis et al., 2020a) and add them as an additional attribute to reduce token consumption.

Ladder Retrieval. Given a question q, we aim to retrieve the most relevant records from $file(P)$. We first utilize an LLM to decompose q into p^* (the primary political actor) and q^* (the question focus conveying the essential semantic content needed to infer an appropriate response (Moldovan et al., 1999)). For instance, in "What vote will the {political actor} cast on {a description of a bill}?", q^* corresponds to the bill description. This decomposition is formalized as $q \xrightarrow{\text{LLM}} (p^*, q^*)$. Once p^* is obtained, we match it to its corresponding political actor p_s and retrieve a set of highly relevant records from $file(p_s)$. Since each r_p contains multiple attributes, traditional embedding-based retrieval methods (Karpukhin et al., 2020; Khattab and Zaharia, 2020) often struggle to effectively exploit such structural attribute information and may introduce noise. Inspired by (Siddiquie et al., 2011; Macdonald and Tonellotto, 2021), we propose ladder retrieval, an iterative retrieval strategy formalized as follows:

$$R^{(i)} = LR(R^{(i-1)}, q^*, k_i). \quad (1)$$

This iterative process begins by initializing the record set $R^{(0)}$ as $file(p_s)$. During the i-th iteration, the retrieval function LR selects k_i records from $R^{(i-1)}$ to generate $R^{(i)}$. The selection is based on the semantic similarity between q^* and the i-th attribute of each record in $R^{(i-1)}$. After n iterations, this process yields $R^{(n)}$, the top k_n records most relevant to q.

> You will be provided with background information. Your task is to extract stance-relevant details that might influence {political_actor}'s vote on the bill: {bill_content}. Your goal is not to answer the question directly, but to identify and list the specific details that are most relevant to understanding how {political_actor} might vote. Keep it concise. Take your time to consider.
>
> Background: $R^{(n)}$

Figure 2: The prompt template of stance inference for distillation.

Stance Inference. Factual political records provide indirect clues rather than explicit answers. We therefore extract their implicit preferences as concise, readable stance evidence. Compared with raw facts, such evidence is more issue-aligned and exposes cues such as behavioral consistency and ideological orientation, providing a more targeted basis for LLM reasoning. To reduce overhead, we distill a large teacher model into a SLM via task-specific knowledge distillation (Gou et al., 2021). We construct political questions from behavior profiles by casting each record as a question, then retrieve the associated record set $R^{(n)}$ using ladder retrieval (Formula (1)). We prompt the teacher to extract stance-relevant information from $R^{(n)}$, yielding high-quality task data (template in Figure 2). During training, the SLM takes the same inputs as the teacher and is optimized with a cross-entropy loss between its logits and the teacher's tokenized outputs, aligning its stance extraction behavior. Thus, q and $R^{(n)}$ can be input into the trained SLM via a predefined template as follows:

$$R^{(n)}, q \xrightarrow{\text{SLM}} \mathcal{S}, \tag{2}$$

where $\mathcal{S}$ is the question-related stance (detailed in Appendix).

2.2 High-Order Structure Signals Acquisition

Actor Interaction Graph. Although political behavior profiles are independent, their actor-centered semi-structured format naturally supports an interaction graph, where political actors are linked by shared records. This allows PSL to mine structural influence implicit in the data. Specifically, we construct a node for each actor and record. For $r_p \in file(p)$, we construct an edge between the actor p and the record r_p. We represent the resulting interaction graph as $\mathcal{G}_\mathcal{I} = \{(p, e_{pr}, r)\} \subseteq P \times E \times R$, where P is the actor set, R is the record set, and E is the set of edges representing interactions between actors and records. For each edge $e_{pr} \in E$ connecting p and r, its weight is defined as follows: for a voting record, the weight is set to 1 if p voted "*Yea*" on r, and -1 otherwise. For a diplomatic record, the weight is set to 1, indicating the existence of an interaction between p and r. In this case, the absence of an interaction is denoted by the non-existence of e_{pr} rather than being assigned a weight of -1. For example, $p_1 \xrightarrow{e_{p_1 r_5}} r_5 \xleftarrow{e_{p_3 r_5}} p_3$ shows how a shared record connects two actors.

Neighbor-Aware Actor Representation Construction. To obtain actor representations that capture structural influence, we propagate information over $\mathcal{G}_I$. Similar to prior work (He et al., 2020), our main goal is to extract structure signals from $\mathcal{G}_I$. We therefore adopt a simple weighted-sum aggregator that iteratively smooths node embeddings over the graph:

$$\begin{aligned} \mathbf{e}_p^{(i+1)} &= \sum_{r \in \mathcal{N}_p} e_{pr} \frac{1}{\sqrt{|\mathcal{N}_p||\mathcal{N}_r|}} \mathbf{e}_r^{(i)}, \\ \mathbf{e}_r^{(i+1)} &= \sum_{p \in \mathcal{N}_r} e_{pr} \frac{1}{\sqrt{|\mathcal{N}_p||\mathcal{N}_r|}} \mathbf{e}_p^{(i)}. \end{aligned} \tag{3}$$

Here, the symmetric normalization term $1/\sqrt{|\mathcal{N}_p||\mathcal{N}_r|}$ follows the standard GCN design (Kipf and Welling, 2017), where $\mathcal{N}_p$ (resp. $\mathcal{N}_r$) denotes the first-hop neighbors of p (resp. r) and $|\mathcal{N}_p|$ (resp. $|\mathcal{N}_r|$) denotes the size of $\mathcal{N}_p$ (resp. $\mathcal{N}_r$). For each record r, we initialize its embedding as $\mathbf{e}_r$ via a frozen pre-trained embedding model (Reimers and Gurevych, 2019). The initial embeddings of political actors can be calculated by aggregating their neighbors as follows:

$$\begin{aligned} \mathbf{e}_r^{(0)} &= \mathbf{e}_r, \\ \mathbf{e}_p^{(0)} &= \sum_{r \in \mathcal{N}_p} e_{pr} \frac{1}{\sqrt{|\mathcal{N}_p|}\sqrt{|\mathcal{N}_r|}} \mathbf{e}_r. \end{aligned} \tag{4}$$

After l propagation layers, we obtain multi-layer representations of p, denoted as $(\mathbf{e}_p^{(0)}, \mathbf{e}_p^{(1)}, \ldots, \mathbf{e}_p^{(l)})$. Inspired by (Xu et al., 2018), we adopt a layer-aggregation mechanism to generate the political actor's final representation by concatenating the first and last embeddings as $\hat{\mathbf{e}}_p = \mathbf{e}_p^{(0)} || \mathbf{e}_p^{(l)}$, where $||$ denotes the concatenation operation.

2.3 Synergistic Enhancement of LLM Reasoning

To enable joint use of actor-stance and high-order structure signals, PSL introduces two levels of synergy. Within the vector view, actor representations reside outside the LLM's language space and are difficult to use directly. Inspired by prior work showing that LLMs can exploit collaborative vectors (Schafer et al., 2007; Zhang et al., 2024), we innovatively treat the actor as a user-like entity and the question focus as an item-like entity, and model their interaction through the actor and question-focus embeddings. The co-embedding module

does more than reduce dimensionality into a shared latent space. By modeling actor–question interactions, it makes actor representations question-conditioned and acts as an implicit structural filter, highlighting relevant neighbors while suppressing structural noise. Specifically, since the dimension of the corresponding actor embedding $\hat{\mathbf{e}}_{p_s}$ is twice that of the question focus embedding $\mathbf{e}_{q^*}$, we first align them by duplicating and concatenating $\mathbf{e}_{q^*}$ with itself, yielding $\hat{\mathbf{e}}_{q^*} = \mathbf{e}_{q^*}||\mathbf{e}_{q^*}$ to ensure dimensional consistency. Here, q^* and p_s are obtained from § 2.1, and $\mathbf{e}_{q^*}$ and $\mathbf{e}_{p_s}$ are their embeddings, respectively. The computation of $\hat{\mathbf{e}}_{p_s}$ is provided in § 2.2. Then, we feed $\hat{\mathbf{e}}_{p_s}$ and $\hat{\mathbf{e}}_{q^*}$ into the co-embedding module (*i.e.*, a trained MLP) separately to generate their collaborative representation:

$$\hat{\mathbf{e}}_{p_s}, \hat{\mathbf{e}}_{q^*} \xrightarrow{\text{MLP}} \hat{\mathbf{e}}^*_{p_s}, \hat{\mathbf{e}}^*_{q^*}. \tag{5}$$

This process enables the LLM to more efficiently leverage their collaborative relationship to capture and utilize the high-order structure signals encoded in the actor representation.

Next, to achieve synergy across the semantic and vector views, we integrate political actor stances with collaborative embeddings to support the LLM's reasoning. This is achieved by populating a predefined instruction template T with both explicit and implicit signals, and then injecting the filled template into the LLM to generate final answers that are grounded in the associated external knowledge:

$$T(q, \mathcal{S}, \hat{\mathbf{e}}^*_{p_s}, \hat{\mathbf{e}}^*_{q^*}) \xrightarrow{\text{LLM}_{\Phi+\Phi'}} Answer, \tag{6}$$

where $\text{LLM}_{\Phi+\Phi'}$ denotes the LLM slightly fine-tuned with LoRA (Hu et al., 2022). The tuning process will be detailed in the next section.

2.4 Learning

MLP Training. To co-embed neighbor-aware political actor representations with question representations, we leverage the records of political actors as q^* to train the MLP defined by Formula (5). We design an objective function based on the Bayesian personalized ranking loss, which encourages the model to assign higher prediction scores to observed (positive) records than to unobserved (negative) ones:

$$\mathcal{L} = -\sum_{p \in P} \sum_{r^+ \in \mathcal{N}_p^+} \sum_{r^- \in \mathcal{N}_p^-} \ln \sigma(y_{pr^+} - y_{pr^-}) + \lambda \|\Theta\|^2, \tag{7}$$

where $y_{pr} = (\hat{\mathbf{e}}^*_p)^T \hat{\mathbf{e}}^*_r$ denotes the inner product between $\hat{\mathbf{e}}^*_p$ and $\hat{\mathbf{e}}^*_r$, λ controls the L_2 regularization strength, Θ denotes the parameters of the MLP, $\mathcal{N}_p^+$ denotes the set of positive (observed) records, referring to those records for which the weight of e_{pr} is 1, $\mathcal{N}_p^-$ denotes the set of negative (unobserved) records, including explicit negative voting records (the weight of e_{pr} is -1) and unobserved diplomatic records (e_{pr} does not exist). More details are provided in Appendix.

LoRA Tuning. We randomly select 800 pairs of $\hat{\mathbf{e}}_p$ and $\hat{\mathbf{e}}_r$, where $\hat{\mathbf{e}}_r$ is uniformly sampled from $\mathcal{N}_p^+$ and $\mathcal{N}_p^-$. These pairs are processed by the trained MLP to produce collaborative embeddings $\hat{\mathbf{e}}^*_p$ and $\hat{\mathbf{e}}^*_r$. Based on collaborative embeddings, we construct instructional samples via the natural language concatenation of $\hat{\mathbf{e}}^*_p$ and $\hat{\mathbf{e}}^*_r$ as input, and the relation between p and r, that is, the corresponding relation label e_{pr} as output. These instructional samples are then used to fine-tune the LLM using LoRA (detailed in Appendix).

3 Experiments

3.1 Experimental Setting

Datasets. We conduct experiments on three available datasets (Mou et al., 2024) covering different political scenarios: (1) RCVP; (2) ICEWS; (3) StaId. The source code and datasets used in our paper are publicly available[4].

Evaluation Metrics. Following (Mou et al., 2024), we adopt the macro F1 score as the evaluation metric for the binary classification tasks (*i.e.*, RCVP and StaId). For the multiple-choice task in ICEWS, we report accuracy.

Setting Details. To assess the effectiveness of PSL, we employ Llama-3.1-8B-Instruct (Grattafiori et al., 2024), Mistral-7B-Instruct (Jiang et al., 2023), Deepseek-7B-Chat (Bi et al., 2024), and GPT-3.5-Turbo[5] as backbone models, with all experiments conducted under this default configuration unless otherwise noted. In the distillation stage, Flan-T5-Small (Chung et al., 2024) is used as the student SLM, while GPT-4o-mini[6] serves as the teacher model. The number of iterations in the ladder retrieval process, retrieved candidates k_n, and propagation layers l are set to 3, 5, and 2, respectively. The data used to train the MLP is entirely from the profile set. The net-

[4]https://github.com/zhouzihan-liu/PSL
[5]https://platform.openai.com/docs/models/gpt-3.5-turbo
[6]https://platform.openai.com/docs/models/gpt-4o-mini

Method	*RCVP*			*ICEWS*			*StaId*		
	Llama-8B	*Mistral-7B*	*Deepseek-7B*	*Llama-8B*	*Mistral-7B*	*Deepseek-7B*	*Llama-8B*	*Mistral-7B*	*Deepseek-7B*
Vanilla	27.36	27.33	28.03	18.42	19.18	19.67	35.00	<u>33.36</u>	29.59
GKP (ACL'22)	23.07	26.29	24.28	7.09	14.37	9.65	<u>38.26</u>	33.04	34.76
RECITE (ICLR'23)	25.08	23.38	14.90	14.89	16.95	11.81	31.88	29.47	27.49
LangChain	31.10	35.52	24.37	8.06	18.06	19.06	27.71	29.51	30.07
InstructRAG (ICLR'25)	29.84	21.64	26.81	11.57	7.71	18.21	31.53	28.30	32.97
KAPING (ACL'23)	23.15	23.96	23.59	20.52	<u>22.52</u>	15.36	36.20	30.28	34.22
MindMap$_{route}$ (ACL'24)	20.57	25.36	22.35	13.54	19.23	14.89	34.98	32.31	31.76
MindMap$_{lang}$ (ACL'24)	21.37	24.19	21.91	12.69	19.04	11.61	35.69	31.51	32.51
MindMap (ACL'24)	18.66	23.02	21.54	12.70	22.01	21.70	34.72	31.97	31.51
PAA (AAAI'25)	<u>44.35</u>	<u>41.78</u>	<u>42.35</u>	<u>31.35</u>	18.99	<u>24.53</u>	36.34	27.41	<u>39.89</u>
PSL	**55.92**	**50.30**	**43.59**	**57.81**	**54.69**	**53.42**	**48.50**	**48.83**	**50.58**

Table 1: Experimental results of PSL with baselines. The best results are highlighted in bold, and the second best results are underlined.

work architecture of the MLP is configured as $1536 \rightarrow 1024 \rightarrow 512 \rightarrow 256 \rightarrow 64 \rightarrow 32$. The regularization coefficient λ is set to 1×10^{-4}, and dropout is applied during training to improve generalization. Following (Zheng et al., 2024), we fine-tune Llama using 800 samples for 3 epochs, with all other hyperparameters kept at their default settings.

3.2 Effectiveness Study

We compare PSL with the following 12 methods. Vanilla, *i.e.*, directly inputs questions into the LLM without any additional knowledge. GKP (Liu et al., 2022) prompts the LLM to generate knowledge relevant to the question and then appends these generated statements into the prompt to improve reasoning. RECITE (Sun et al., 2023) prompts the LLM to "recite" self-sampled memory-like passages before answering, leveraging internalized knowledge for better QA performance. LangChain (Chase, 2022) retrieves political documents based on semantic similarity and uses them as prompts to improve answer generation. InstructRAG (Wei et al., 2025) prompts the LLM to generate reasoning steps before answering, helping it better use the retrieved political documents. KAPING (Baek et al., 2023) retrieves political knowledge from KG based on semantic similarity and formats them as triples in prompts. MindMap (Wen et al., 2024) generates both the answer and a reasoning path in a tree-like format for interpretability. MindMap$_{route}$ structures retrieved triples into paths, such as "climate policy debate $\rightarrow$ related politicians $\rightarrow$ John Smith, Emily Carter", and uses them to guide reasoning. MindMap$_{lang}$ converts the structured knowledge paths into natural language narratives, which are then used to prompt the LLM. PEG$_{exp_sum}$ (Mou et al., 2024) retrieves relevant political triples and summarizes them using the LLM to facilitate downstream reasoning. PEG$_{exp_GTR}$ (Mou et al., 2024) clusters and summarizes selected triples via LLMs for final answer generation. PAA (Li et al., 2025) utilizes political actor profiles to guide LLMs in simulating legislative behavior for roll-call vote prediction. For document-enhanced methods (*i.e.*, LangChain and InstructRAG), we treat each record in profiles as an independent document. For KG-enhanced methods (*i.e.*, KAPING, MindMap$_{route}$, MindMap$_{lang}$, MindMap, PEG$_{exp_sum}$, and PEG$_{exp_GTR}$), the political KG MVPKG (Mou et al., 2024) is uniformly adopted as the external knowledge source. For the agent-based method PAA, we follow the original approach and extract 20 records per political actor from our constructed profiles. Note that the raw data used to construct MVPKG fully encompasses the profiles we construct. To report the comparative results with PEG, since its core implementation has not been publicly released, we present the performance metrics as reported in the original paper. The results are summarized in Table 2.

From the results shown in Table 1, we can see that PSL outperforms all baselines on three datasets with different LLMs. Models without external knowledge (*i.e.*, GKP and RECITE) usually fail to surpass Vanilla, as LLMs struggle to generate accurate knowledge about political facts or future events. Document-enhanced methods perform well only over RCVP, with limited effectiveness elsewhere, due to challenges in accurately retrieving relevant information and the interference caused by direct text injection in complex reasoning over ICEWS and StaId. In contrast, KG-enhanced methods show stronger and more consistent performance. KAPING remains stable across three datasets, while MindMap and its variants, relying on entity-path retrieval, are less effective in future-oriented scenarios. PAA achieves notable performance over

Method	*RCVP*	*ICEWS*	*StaId*
Vanilla	34.11	19.40	33.24
GKP (ACL'22)	20.43	15.40	45.32
RECITE (ICLR'23)	15.79	21.40	35.57
KAPING (ACL'23)	37.83	20.80	36.99
MindMap$_{route}$ (ACL'24)	32.60	28.00	35.03
MindMap$_{lang}$ (ACL'24)	38.57	28.40	42.19
MindMap (ACL'24)	38.72	25.60	23.38
PEG$_{exp_sum}$ (WWW'24)	40.62	26.40	42.12
PEG$_{exp_GTR}$ (WWW'24)	<u>41.21</u>	<u>28.60</u>	<u>46.21</u>
PSL	**50.23**	**45.89**	**49.90**

Table 2: Experimental results of PSL with all baselines under GPT-3.5. The best results are highlighted in bold, and the second best results are underlined. Results are taken from (Mou et al., 2024), except for PSL.

RCVP. Although both PEG and PAA are designed to enhance LLMs in political tasks using external knowledge similar in content to that employed by PSL, PSL consistently outperforms them in all experimental setups, which may be attributed to the fact that PSL can mine political knowledge from semi-structured data to enhance the LLM. Furthermore, PSL also demonstrates competitive time efficiency, as detailed in Appendix.

3.3 Ablation Study

We validate PSL's components by removing actor stance (*w/o Stance*) or high-order structure (*w/o Structure*) on Llama-8B. Note that *w/o Structure* eliminates collaborative embeddings, restricting inputs to stance and question text. Results in Table 3 show: (1) *w/o Stance* causes significant drops across all datasets, validating that retrieved records indeed capture critical issue-specific positions; (2) *w/o Structure* degrades performance, notably on RCVP and ICEWS , confirming that high-order signals capture group-level behavioral dependencies essential for voting and diplomatic prediction.

3.4 Parameter Study

We examine the impact of retrieved records and propagation layers using Llama-8B (Figure 3). (1) Retrieved Records: Performance improves initially, peaking at 5 records for RCVP/ICEWS and 15 for StaId, before declining due to noise from irrelevant content. StaId shows a wider growth range, suggesting a need for richer context. Notably, PSL surpasses baselines across all settings. (2) Propagation Layers: Performance peaks at layer 2 (RCVP/ICEWS) and layer 3 (StaId). This confirms that limited propagation captures essential structure signals, whereas deeper layers lead to over-smoothing and reduced discriminative power.

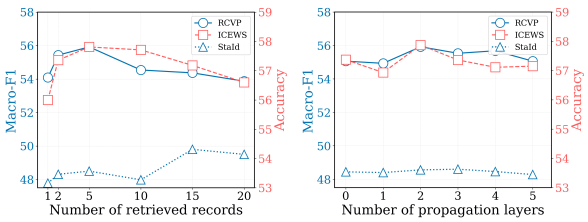
Ablations	*RCVP*	*ICEWS*	*StaId*
PSL	**55.92**	**57.81**	**48.50**
w/o Actor Stance Acquisition	46.15	39.07	46.17
w/o High-Order Structure Signals Acquisition	39.35	49.02	46.35

Table 3: The results of ablation study for PSL.

4 Further Analysis

Effect Analysis of Different Components of Actor Stance Acquisition. We investigate the impact of retrieval and stance inference by evaluating PSL variants (without stance inference) using TF-IDF (Aizawa, 2003), BM25 (Robertson et al., 2009), Embedding (Reimers and Gurevych, 2019), and our Ladder Retrieval (PSL$_{LR}$). As shown in Table 4, all variants surpass the baselines and the "w/o Stance" ablation, demonstrating that profile records provide critical context for prediction. Specifically, lexical methods (BM25, TF-IDF) yield lower scores due to their reliance on lexical overlap, failing to capture semantic or structural dependencies. Conversely, PSL$_{LR}$ achieves the best performance among retrieval variants. This validates LR's effectiveness in leveraging hierarchical structural information and handling missing data through score inheritance. Ultimately, the full PSL framework outperforms all variants, confirming that transforming isolated records into targeted stance information reduces noise and boosts overall performance.

Effect Analysis of The Presentation Form for High-Order Structure Signals. We investigate representing high-order structure as text by retrieving records from similar neighbors. For each actor, we rank its directly connected neighbors by this similarity and select the top ones. Similarity between two actors is computed from their shared record links: for each record connected to both, we add 1 if their edge weights match (both +1 or both -1) and subtract 1 if they differ. Figure 4 shows the

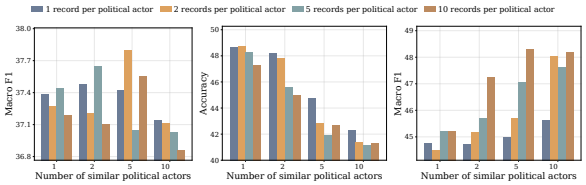


(a) With varied parameter of the retrieved record number. (b) With varied parameter of the propagation layer number.

Figure 3: Parameter study on the RCVP, ICEWS, and StaId datasets.

Variant	*RCVP*	*ICEWS*	*StaId*
PSL	**55.92**	**57.81**	**48.50**
PSL$_{TF-IDF}$ w/o SI	52.07	56.70	46.45
PSL$_{BM25}$ w/o SI	51.87	57.22	45.94
PSL$_{Embedding}$ w/o SI	52.96	57.16	46.69
PSL$_{LR}$ w/o SI	53.81	57.34	47.10

Table 4: Performance of different variants of PSL. SI denotes the procedure of stance inference.



(a) RCVP (b) ICEWS (c) StaId

Figure 4: Performance of PSL using different numbers of political actors' neighbors and their records in the form of text.

impacts. RCVP benefits from distributing retrieval across more actors, as top-ranked neighbors provide the most valuable signals. Conversely, ICEWS performance declines with increased context due to noise from indirect records, whereas StaId improves with retrieval volume to compensate for missing statement profiles. Crucially, all text-based settings consistently underperform PSL. This corroborates our assertion that articulating complex relations in natural language leads to information overload, validating the necessity of PSL's collaborative vector modeling.

5 Related Work

LLM Augmentation via External Knowledge. External knowledge augmentation grounds LLMs in resources beyond their parameters. Document-based methods, such as REALM (Guu et al., 2020), RAG (Lewis et al., 2020b), and RePlug (Shi et al., 2024), retrieve relevant passages, whereas KG-based methods inject compact factual triples. KAPING (Baek et al., 2023) and DiagLink (Zhou et al., 2026) uses KG triples for question answering; GPT4Graph (Guo et al., 2023) and NLGraph (Wang et al., 2023) study graph reasoning with LLMs; and ToG (Sun et al., 2024) and MindMap (Wen et al., 2024) construct reasoning paths over graph structure. These methods are mainly designed for settings where external knowledge can provide answers or explicit reasoning paths. Predictive political QA is different: external resources record past behavior rather than the target behavior itself. Recent work has adapted external augmentation to political reasoning. PEG (Mou et al., 2024) injects retrieved triples from a political KG, whereas PAA (Li et al., 2025) simulates political actors from textual profiles. Both improve political reasoning, but their evidence forms are limited for prediction: KG triples lose fine-grained context, and textual profiles weakly capture implicit group influence. We instead derive stance evidence and structure signals from semi-structured political records.

Political Actor Modeling. Political actor modeling is central to computational political science. Classical approaches estimate actors' ideological positions from roll-call votes, most notably through ideal point models (Clinton et al., 2004), later extended with legislative text (Gerrish and Blei, 2011; Kraft et al., 2016). Graph-based methods further model political relations through co-sponsorship networks (Yang et al., 2020), contribution networks (Davoodi et al., 2020), and donation networks (Davoodi et al., 2022). Recent work enriches actor representations with external knowledge, including media coverage (Feng et al., 2021), social media statements (Mou et al., 2021), Wikipedia, and policy-research materials (Feng et al., 2022). Despite their effectiveness, the heterogeneity of these data sources results in high collection costs. In this paper, we construct political behavior profiles for political actors using open-access and continuously updated sources and inject actors' stances and their complex political relationships into LLMs.

6 Conclusion

We propose PSL, a framework that transforms semi-structured political records into inference-oriented evidence for predictive political QA. PSL enables textual semantics and graph-structural features to jointly strengthen LLM reasoning. This dual-view design, grounded in semi-structured data, provides a more prediction-oriented evidence representation. Moreover, compared with KG triples, PSL better preserves contextual detail; compared with text-profile simulation, it more fully models structural relations among actors. Experiments on multiple real-world datasets and LLMs show that PSL consistently outperforms existing augmentation methods. More broadly, our results provide initial evidence that inference-oriented evidence beyond knowledge-based evidence is effective for predictive political QA. We will release the code and

model checkpoints to facilitate further research.

Limitations

Data scope. This paper validates PSL in U.S. political settings, using public records such as legislative votes, bill information, and diplomatic events to construct semi-structured profiles and an interaction graph. This setting provides a clear experimental basis for testing PSL's core design: extracting actor stance signals and high-order structure signals from commonly available political records to enhance predictive political QA. However, there remains room to broaden the data scope. On the one hand, future work could apply PSL to more countries, political systems, and language contexts to assess its generalizability across political environments. On the other hand, existing records could also be extended to richer sources, such as co-sponsorship, committee activities, public statements, campaign donations, and policy texts. These extensions do not require changes to PSL's basic design, but could provide more comprehensive behavioral evidence for stance extraction and structure modeling.

Signal fusion. This paper designs two complementary signals to enhance predictive political QA. Experiments and ablation studies show that these signals provide effective reasoning evidence for LLMs from different perspectives. However, different questions may rely on the two signals to different degrees. Some questions may depend more on an actor's historical stance on related issues, whereas others may rely more on group-level behavioral patterns among political actors. Future work could explore adaptive fusion mechanisms that dynamically adjust the weights of the two signals according to the question, actor, or political setting, enabling finer-grained modeling of evidence needs across predictive tasks.

Acknowledgments

The work was partially supported by the National Key Research and Development Program of China (No. 2024YFF0617702); the National Natural Science Foundation of China (Nos. U22A2025, 62402097, 62232007, and U23A20309); the Joint Funds of the Natural Science Foundation of Liaoning Province (No. 2023-BSBA-132); the 111 Project (No. B16009); the Ant Group Research Program (No. 2025021900003); and the Fundamental Research Funds for the Central Universities (No. N2417007)

References

Akiko Aizawa. 2003. An information-theoretic perspective of tf–idf measures. *IPM*, 39(1):45–65.

Jinheon Baek, Alham Fikri Aji, and Amir Saffari. 2023. Knowledge-augmented language model prompting for zero-shot knowledge graph question answering. In *NLRSE*, pages 78–106.

Xiao Bi, Deli Chen, Guanting Chen, Shanhuang Chen, Damai Dai, Chengqi Deng, Hongyuan Ding, Kai Dong, Qiushi E, and 1 others. 2024. Deepseek llm: Scaling open-source language models with longtermism. *arXiv*.

Elizabeth Boschee, Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, James Starz, and Michael Ward. 2015. ICEWS Coded Event Data.

Michael Burnham. 2025. Stance detection: a practical guide to classifying political beliefs in text. *Political Science Research and Methods*, 13(3):611–628.

Harrison Chase. 2022. LangChain.

Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, and 1 others. 2024. Scaling instruction-finetuned language models. *JMLR*, 25(70):1–53.

Joshua Clinton, Simon Jackman, and Douglas Rivers. 2004. The statistical analysis of roll call data. *American Political Science Review*, 98(2):355–370.

Florin Cuconasu, Giovanni Trappolini, Federico Siciliano, Simone Filice, Cesare Campagnano, Yoelle Maarek, Nicola Tonellotto, and Fabrizio Silvestri. 2024. The power of noise: Redefining retrieval for RAG systems. In *SIGIR*, pages 719–729.

Maryam Davoodi, Eric Waltenburg, and Dan Goldwasser. 2020. Understanding the language of political agreement and disagreement in legislative texts. In *ACL*, pages 5358–5368.

Maryam Davoodi, Eric Waltenburg, and Dan Goldwasser. 2022. Modeling U.S. state-level policies by extracting winners and losers from legislative texts. In *ACL*, pages 270–284.

Shangbin Feng, Zilong Chen, Qingyao Li, and Minnan Luo. 2021. Kgap: Knowledge graph augmented political perspective detection in news media. *CoRR*, abs/2108.03861.

Shangbin Feng, Zhaoxuan Tan, Zilong Chen, Ningnan Wang, Peisheng Yu, Qinghua Zheng, Xiaojun Chang, and Minnan Luo. 2022. PAR: Political actor representation learning with social context and expert knowledge. In *EMNLP*, pages 12022–12036.

Divyanshi Galla and James Burke. 2018. Predicting social unrest using GDELT. In *MLDM*, pages 103–116.

Sean Gerrish and David M. Blei. 2011. Predicting legislative roll calls from text. In *ICML*, pages 489–496.

Jianping Gou, Baosheng Yu, Stephen J Maybank, and Dacheng Tao. 2021. Knowledge distillation: A survey. *IJCV*, 129(6):1789–1819.

Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, and et al. 2024. The llama 3 herd of models. *CoRR*, abs/2407.21783.

Jiayan Guo, Lun Du, and Hengyu Liu. 2023. Gpt4graph: Can large language models understand graph structured data ? an empirical evaluation and benchmarking. *CoRR*, abs/2305.15066.

Kelvin Guu, Kenton Lee, Zora Tung, Panupong Pasupat, and Ming-Wei Chang. 2020. Retrieval augmented language model pre-training. In *ICML*, pages 3929–3938.

Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. 2020. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *SIGIR*, pages 639–648.

Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. LoRA: Low-rank adaptation of large language models. In *ICLR*, pages 12513–12525.

Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Chaplot Devendra, Guillaume Lample, Kevin Leach, Pierre Stock, Teven Le Scao, and 1 others. 2023. Mistral 7b. *arXiv*.

Vladimir Karpukhin, Barlas Oguz, Sewon Min, Patrick SH Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih. 2020. Dense passage retrieval for open-domain question answering. In *EMNLP*, pages 6769–6781.

Omar Khattab and Matei Zaharia. 2020. Colbert: Efficient and effective passage search via contextualized late interaction over BERT. In *SIGIR*, pages 39–48.

Thomas N. Kipf and Max Welling. 2017. Semi-supervised classification with graph convolutional networks. In *ICLR*, pages 2713–2726.

Peter Kraft, Hirsh Jain, and Alexander M. Rush. 2016. An embedding model for predicting roll-call votes. In *EMNLP*, pages 2066–2070.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020a. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *ACL*, pages 7871–7880.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich K¨uttler, Mike Lewis, Wen-tau Yih, Tim Rockt¨aschel, Sebastian Riedel, and Douwe Kiela. 2020b. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *NeurIPS*, pages 9459–9474.

Hao Li, Ruoyuan Gong, and Hao Jiang. 2025. Political actor agent: Simulating legislative system for roll call votes prediction with large language models. In *AAAI*, pages 388–396.

Jiacheng Liu, Alisa Liu, Ximing Lu, Sean Welleck, Peter West, Ronan Le Bras, Yejin Choi, and Hannaneh Hajishirzi. 2022. Generated knowledge prompting for commonsense reasoning. In *ACL*, pages 3154–3169.

Craig Macdonald and Nicola Tonellotto. 2021. On approximate nearest neighbour selection for multi-stage dense retrieval. In *CIKM*, pages 3318–3322.

Dan I. Moldovan, Sanda M. Harabagiu, Marius Pasca, Rada Mihalcea, Richard Goodrum, Roxana Girju, and Vasile Rus. 1999. LASSO: A tool for surfing the answer net. In *TREC*, pages 65–73.

Xinyi Mou, Zejun Li, Hanjia Lyu, Jiebo Luo, and Zhongyu Wei. 2024. Unifying local and global knowledge: Empowering large language models as political experts with knowledge graphs. In *WWW*, pages 2603–2614.

Xinyi Mou, Zhongyu Wei, Lei Chen, Shangyi Ning, Yancheng He, Changjian Jiang, and Xuanjing Huang. 2021. Align voting behavior with public statements for legislator representation learning. In *ACL/IJCNLP*, pages 1236–1246.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *EMNLP-IJCNLP*, pages 3980–3990.

Stephen Robertson, Hugo Zaragoza, and 1 others. 2009. The probabilistic relevance framework: Bm25 and beyond. *FTIR*, 3(4):333–389.

Pranab Sahoo, Prabhash Meharia, Akash Ghosh, Sriparna Saha, Vinija Jain, and Aman Chadha. 2024. A comprehensive survey of hallucination in large language, image, video and audio foundation models. In *EMNLP*, pages 11709–11724.

J. Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen. 2007. Collaborative filtering recommender systems. *The Adaptive Web*, pages 291–324.

Weijia Shi, Sewon Min, Michihiro Yasunaga, Minjoon Seo, Richard James, Mike Lewis, Luke Zettlemoyer, and Wen-tau Yih. 2024. REPLUG: retrieval-augmented black-box language models. In *NAACL*, pages 8371–8384.

Behjat Siddiquie, Rogério Schmidt Feris, and Larry S. Davis. 2011. Image ranking and retrieval based on multi-attribute queries. In *CVPR*, pages 801–808.

Sophie Sowden, Sofia Koletsi, Eva Lymberopoulos, Elisabeta Militaru, Caroline Catmur, and Geoffrey Bird. 2018. Quantifying compliance and acceptance through public and private social conformity. *Consciousness and Cognition*, 65:359–367.

Jiashuo Sun, Chengjin Xu, Lumingyuan Tang, Saizhuo Wang, Chen Lin, Yeyun Gong, Lionel M. Ni, Heung-Yeung Shum, and Jian Guo. 2024. Think-on-graph: Deep and responsible reasoning of large language model on knowledge graph. In *ICLR*, pages 14199–14229.

Zhiqing Sun, Xuezhi Wang, Yi Tay, Yiming Yang, and Denny Zhou. 2023. Recitation-augmented language models. In *ICLR*, pages 31171–31193.

Heng Wang, Shangbin Feng, Tianxing He, Zhaoxuan Tan, Xiaochuang Han, and Yulia Tsvetkov. 2023. Can language models solve graph problems in natural language? In *NeurIPS*, pages 30840–30861.

Jingjin Wang and Jiawei Han. 2025. PropRAG: Guiding retrieval with beam search over proposition paths. In *EMNLP*, pages 6223–6238.

Zhepei Wei, Wei-Lin Chen, and Yu Meng. 2025. InstructRAG: Instructing retrieval-augmented generation via self-synthesized rationales. In *ICLR*, pages 66510–66533.

Yilin Wen, Zifeng Wang, and Jimeng Sun. 2024. Mindmap: Knowledge graph prompting sparks graph of thoughts in large language models. In *ACL*, pages 10370–10388.

Keyulu Xu, Chengtao Li, Yonglong Tian, Tomohiro Sonobe, Ken-ichi Kawarabayashi, and Stefanie Jegelka. 2018. Representation learning on graphs with jumping knowledge networks. In *ICML*, pages 5449–5458.

Yuqiao Yang, Xiaoqiang Lin, Geng Lin, Zengfeng Huang, Changjian Jiang, and Zhongyu Wei. 2020. Joint representation learning of legislator and legislation for roll call prediction. In *IJCAI*, pages 1424–1430.

Yang Zhang, Keqin Bao, Ming Yan, Wenjie Wang, Fuli Feng, and Xiangnan He. 2024. Text-like encoding of collaborative information in large language models for recommendation. In *ACL*, pages 9181–9191.

Yaowei Zheng, Richong Zhang, Junhao Zhang, Yanhan Ye, Zheyan Luo, Zhangchi Feng, and Yongqiang Ma. 2024. Llamafactory: Unified efficient fine-tuning of 100+ language models. In *ACL*, pages 400–410.

Zihan Zhou, Yinan Liu, Yuyang Xie, Bin Wang, Xiaochun Yang, and Zezheng Feng. 2026. Diaglink: A dual-user diagnostic assistance system by synergizing experts with llms and knowledge graphs. In *CHI*, pages 1–28.

A EFFICIENCY STUDY

We conduct experiments on RCVP to evaluate the running efficiency. To minimize the impact of platform differences, all experiments are performed using the GPT-3.5-Turbo API. As shown in Table 5, the intermediate generation process is the primary factor contributing to the efficiency degradation. However, this process is inevitable when interacting with LLMs. Benefiting from the efficient utilization of the SLM, the intermediate generation time of PSL is shorter than that of existing generative approaches (*e.g.*, RECITE, MindMap, and PEG). In future work, we plan to explore more efficient interaction strategies with LLMs.

Method	***RT***	***Process***
Vanilla	1.00	FI
GKP	2.10	IG + FI
RECITE	3.11	IG + FI
LangChain	1.67	R + FI
InstructRAG	1.92	R + IG + FI
KAPING	1.88	R + FI
MindMap$_{route}$	1.95	R + FI
MindMap$_{lang}$	4.99	R + IG + FI
MindMap	13.9	R + IG + FI
PEG$_{exp_sum}$	2.52	R + IG + FI
PEG$_{exp_GTR}$	3.42	R + IG + FI
PAA	1.78	FI
PSL	2.48	R + IG + FI

Table 5: Running efficiency analysis. Relative Time (RT) denotes the ratio of the model's running time using the GPT-3.5-Turbo interface to that of the baseline Vanilla GPT-3.5-Turbo. Process abbreviations: R denotes Retrieval, IG denotes Intermediate Generation, and FI denotes Final Inference.

B EVALUATION

For evaluation, we follow the setting of (Mou et al., 2024). Specifically, we provide multiple options in the prompt and instruct the model to output its choice. To handle cases where the options are not explicitly stated in the output, we employ regular expressions to match the answers effectively.

C MLP TRAINING

In this section, we detail the training procedure for the MLP. The training data is constructed from actor profiles. For each actor, every positive record embedding is paired with a randomly selected negative record embedding to form a training sample, *i.e.*, (actor embedding, positive embedding, negative embedding). Aggregating these samples across all actors yields over 900,000 samples in total. From these, 500,000 samples are randomly selected to construct the dataset, using a random seed of 42, which is then split into training (70%), validation (15%), and test (15%) subsets. The training parameters are summarized in Table 7. The model performance is evaluated by using AUC as the evaluation metric, while both loss and AUC are reported for the test set. Figure 5 presents the training dynamics and final evaluation results.

Parameter	**Value**
Learning Rate	5e-5
Learning Rate Scheduler	Cosine
Epochs	3.0
Warmup Steps	20
Maximum Samples	800
Validation Split Ratio	0.1
Training Batch Size	4
Validation Batch Size	1
Gradient Accumulation Steps	8
Effective Batch Size	32 (4×8)
Maximum Sequence Length	2048
Data Processing Threads	16
Precision Type	BF16

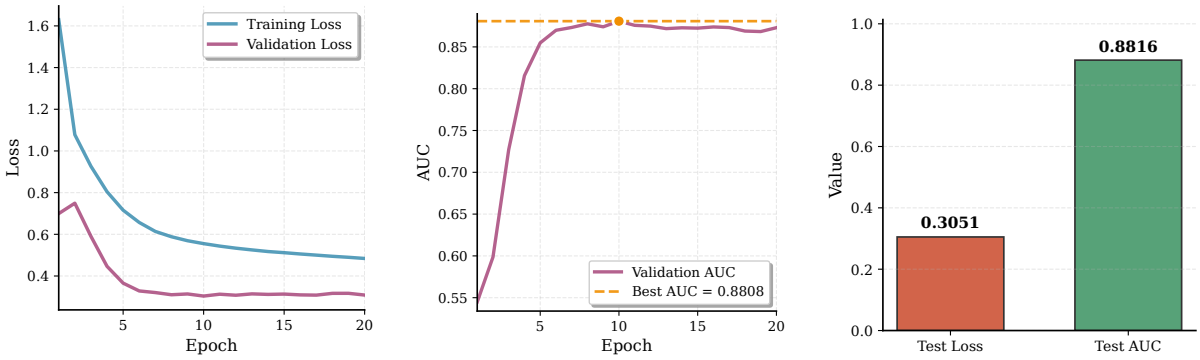
Table 6: LoRA-Tuning parameters.

Parameter	**Value**
Batch Size	40,960
Max Epochs	100
Learning Rate	0.0005
Optimizer	Adam
Weight Decay	0
L2 Regularization	1e-4
Early Stopping Patience	10
Minimum Delta	0

Table 7: MLP training parameters.

D LoRA TUNING

This section details the Low-Rank Adaptation (LoRA) fine-tuning method applied to LLMs. LoRA is a parameter-efficient fine-tuning technique that introduces trainable low-rank decomposition matrices alongside pre-trained weight ma-



(a) Loss curves (b) Validation AUC (c) Test performance

Figure 5: MLP training curves and model performance over epochs.

trices while keeping the original model parameters frozen. This approach significantly reduces the number of trainable parameters, lowering computational and storage costs while effectively adapting to downstream tasks. All experiments were conducted on 4 $\times$ NVIDIA A6000 GPUs. To conserve resources, the total number of training samples was limited to 800. Table 6 summarizes the key parameter settings for LoRA fine-tuning.

E EXAMPLE OF STANCE

Sample stance of Seth Moulton

1. **Previous Votes on Background Checks**: Seth Moulton has consistently voted 'yea' on the Enhanced Background Checks Act of 2019, indicating strong support for measures related to background checks for firearm sales. 2. **Legislative Focus**: Moulton's voting record shows a tendency to support legislation aimed at amending existing laws related to regulatory oversight, including labor relations and background checks. 3. **Consistency in Voting**: The repeated 'yea' votes on the same bill (Enhanced Background Checks Act of 2019) suggest a firm stance on the issue, which could carry over to his vote on H.B. 8 regarding background checks for every firearm sale. 4. **Context of Gun Control Legislation**: Moulton's history of supporting background check legislation may reflect a broader commitment to gun control measures, potentially influencing his decision on related bills. 5. **Political Implications**: As a member of Congress, Moulton may consider the political landscape and public opinion regarding gun control when deciding on votes, especially in light of his previous support for similar bills.

F EXAMPLE OF POLITICAL BEHAVIOR PROFILE

Sample Profile of Lindsey Graham

{
"title": "Pechanga Band of Luiseno Mission Indians Water Rights Settlement Act",
"description": "Pechanga Band of Luiseno Mission Indians Water Rights Settlement Act (Sec. 4) This bill authorizes, ratifies, and confirms the Pechanga Settlement Agreement, entered into by the Pechanga Band of Luiseno Mission Indians, the Rancho California Water District (RCWD), and the United States, except to the extent that the agreement is modified by or conflicts with this bill. (Sec. 5) The bill confirms water rights that must be held in trust by the United States on behalf of the tribe and its allottees. (Allottees are individuals who hold a beneficial real property interest in an Indian allotment that is located within the reservation and held in trust by the United States.) Allotted land is entitled to a just and equitable allocation of water from the water rights for irrigation and domestic purposes. Allottees may lease their land together with any water right. The tribe must enact a Pechanga Water Code that governs the storage, recovery, and use of the water rights, subject to the Department of the Interior's approval. Interior must administer the water rights until the water code is enacted and approved. (Sec. 7) The tribe and the United States (acting as trustee for the tribe and allottees) must waive all claims to water rights within the Santa Margarita River Watershed, except water rights recognized in the Pechanga Settlement Agreement and this bill. The tribe and the United States (acting as trustee for the tribe) waive specified claims against the RCWD. The tribe may waive claims against the United States regarding specified water rights and damages. The waivers in this bill are enforceable on the date Interior publishes specified findings regarding deposits, waivers, and approved agreements. (Sec. 8) Interior must provide the amounts necessary to fulfill the tribe's obligations under specified agreements regarding water infrastructure. (Sec. 9) The bill establishes the Pechanga Settlement Fund. The fund is to be used to carry out this bill. (Sec. 12) If Interior does not publish the findings required for enforcement of the waivers in this bill by April 30, 2021, or an alternative later date agreed to by the tribe and Interior, the provisions of this bill expire, related agreements are void, and funds are rescinded.",
"summary": "The bill confirms that water rights shall be held in trust by the United States for the tribe and its allottees, who are individuals with beneficial interests in trust allotments within the reservation. The tribe must establish a Pechanga Water Code to regulate the storage, recovery, and use of these rights."
},
{
"event_text": "express intent to engage in diplomatic cooperation, including public policy support for Taiwan's international participation, strengthening defense and security ties, and promoting economic exchanges",
"target_name": "taiwan"
}