GenTKG: Generative Forecasting on Temporal Knowledge Graph

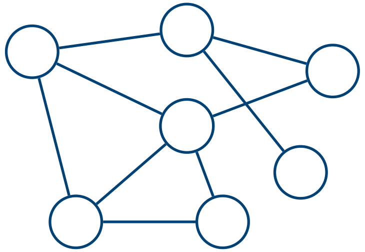
Ruotong Liao, Xu Jia, Yunpu Ma, Volker Tresp

ruotong.liao@outlook.com

16. Dec. 2023

Introduction

- What do we do in this research? $\rightarrow$ Be a fortuneteller and **Forecasting the Future!**
- What do we need? **History information.**
- Where do we get history information from? **Temporal Knowledge Graph** (TKG)



TKG

Quadruples: $(e_s, r, \ e_o, \ t)$

Example:
(Merkel, consult, Obama, 2014/08/09)
$\leftarrow$ (Merkel, discuss by telephone, Obama, 2014/07/22)

TKGF

TKG Forecasting aims to answer test queries:
$(e_s, r, \ ?, \ t)$, given the history facts before t.

Limitations of Current Methods

- Embedding-based methods:
  - Require carefully designed models
  - Lose the semantic aspects of TKGs

  Examples: RE-GCN, xERTE, TANGO, Timetraveler

- Rule-based method:
  - Mines logic rules within the TKG
  - Limited scalability to different datasets

  Example: Tlogic

- LLMs with In-Context Learning :
  - Fail to compete with the above methods

Li, Zixuan & Jin, Xiaolong & Li, Wei & Guan, Saiping & Guo, Jiafeng & Shen, Huawei & Wang, Yuanzhuo & Cheng, Xueqi. (2021). Temporal Knowledge Graph Reasoning Based on Evolutional Representation Learning. 408-417. 10.1145/3404835.3462963.
Zhen Han, Peng Chen, Yunpu Ma, and Volker Tresp. Explainable subgraph reasoning forforecasting on temporal knowledge graphs. In International Conference on Learning Representations, 2020.
Zhen Han, Zifeng Ding, Yunpu Ma, Yujia Gu, and Volker Tresp. Learning neural ordinary equations for forecasting future links on temporal knowledge graphs. In Proceedings of the 2021 conference on empirical methods in natural language processing, pages 8352–8364, 2021.
Haohai Sun, Jialun Zhong, Yunpu Ma, Zhen Han, and Kun He. Timetraveler: Reinforcement learning for temporal knowledge graph forecasting. arXiv preprint arXiv:2109.04101, 2021.
Yushan Liu, Yunpu Ma, Marcel Hildebrandt, Mitchell Joblin, and Volker Tresp. Tlogic: Temporal logical rules for explainable link forecasting on temporal knowledge graphs. In Proceedings of the AAAI conference on artificial intelligence, volume 36, pages 4120–4127, 2022.
Dong-Ho Lee, Kian Ahrabian, Woojeong Jin, Fred Morstatter, and Jay Pujara. Temporal knowledge graph forecasting without knowledge using in-context learning. arXiv preprint arXiv:2305.10613, 2023.

Challenges and Solutions

Can pre-trained LLMs understand structured temporal relational data and replace them as the foundation model for temporal relational forecasting?

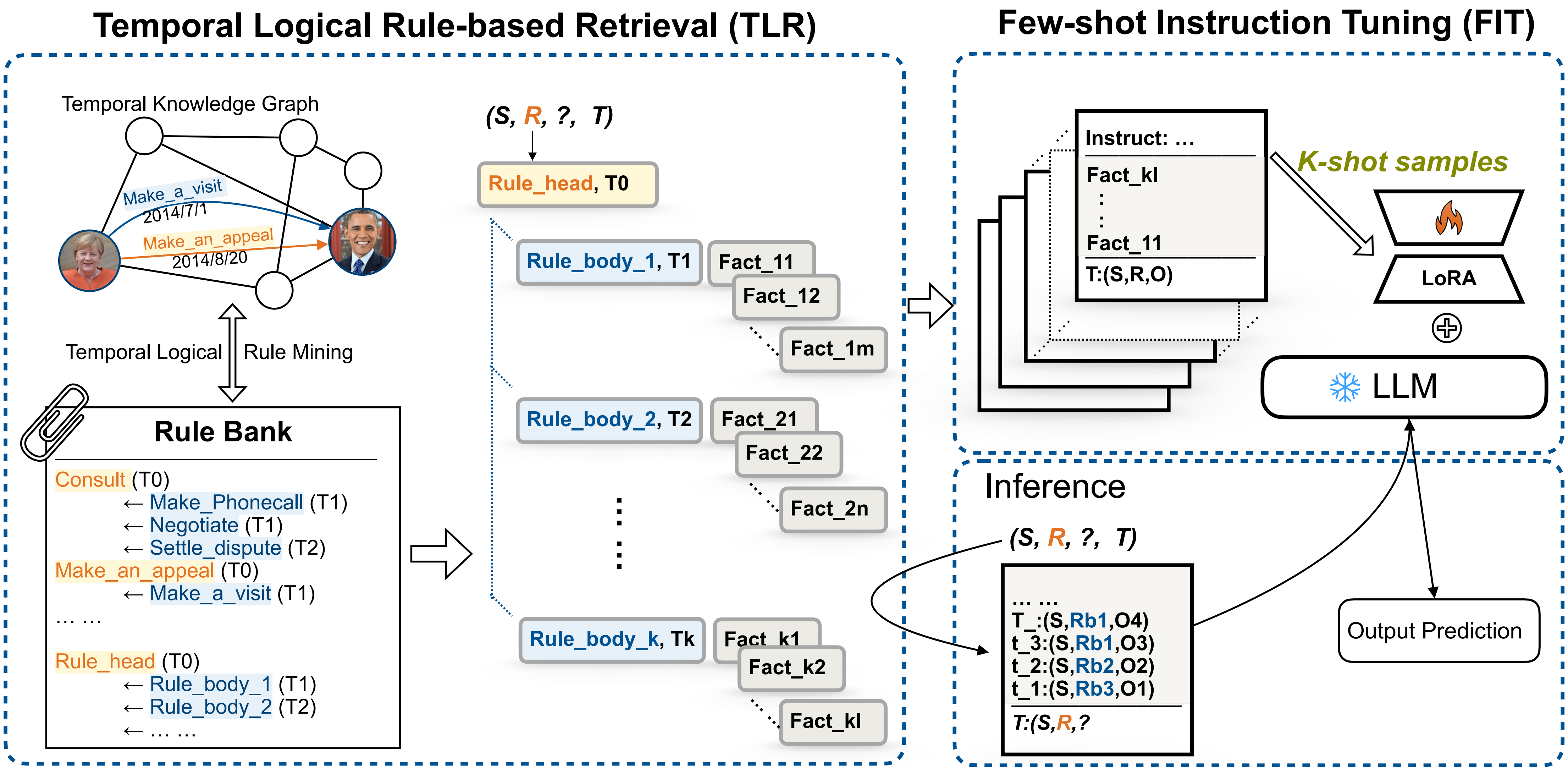
Challenges:
- Modality challenge between data structures
    - TKG: complex temporal multi-relational graph data
    - LLMs: only process sequential natural language expressions

- Computation challenge
    - Hundreds of thousands of quadruples
    - Enormous costs of fine-tuning LLMs

Solutions: a novel retrieval-augmented generation framework, **GenTKG**
- <u>Temporal Logical Rule-based</u> history facts retrieval strategy
    - Enable LLM to comprehend temporal relational data

- <u>Few-shot Parameter-efficient Instruction Tuning</u>
    - As few as 0.0027% training data (16 samples) with drastically low computation

Framework Structure



Performances

Main Results

Datasets	#train	#valid	#test	#entity	#relations	time gap
ICEWS14	74854	8514	7371	7128	230	1 day
ICEWS18	373018	45995	49545	23033	256	1 day
GDELT	79319	9957	9715	5850	238	15 mins
YAGO	220393	28948	22765	10778	23	1 year

Table1. Dataset Statistics

	Datasets	ICEWS14			ICEWS18			GDELT			YAGO		
Method Type	Model	Hits@1	Hits@3	Hits@10	Hits@1	Hits@3	Hits@10	Hits@1	Hits@3	Hits@10	Hits@1	Hits@3	Hits@10
Embedding-based	RE-GCN	0.313	0.473	**0.626**	0.223	0.367	**0.525**	0.084	0.171	0.299	0.468	0.607	0.729
	xERTE	0.330	0.454	0.570	0.209	0.335	0.462	0.112	0.191	0.294	**0.769**	0.787	0.794
	TANGO	0.272	0.408	0.550	0.191	0.318	0.462	0.094	0.189	0.322	0.566	0.651	0.718
	Timetraveler	0.319	0.454	0.575	0.212	0.325	0.439	0.112	0.186	0.285	0.604	0.770	0.831
Rule-based	TLogic	0.332	0.476	0.602	0.204	0.336	0.480	0.113	0.212	0.351	0.638	0.650	0.660
In-Context Learning	GPT-NeoX-20B	0.319	0.439	0.538	0.179	0.297	0.41	0.098	0.165	0.253	0.669	0.787	0.841
	Llama2-7B	0.252	0.427	0.504	0.128	0.272	0.323	0.06	0.164	0.246	0.662	0.760	0.818
GENTKG	GPT-NeoX-20B +TLR	0.35	0.485	0.593	0.205	0.338	0.462	0.156	0.241	0.349	0.681	**0.807**	**0.861**
	Llama2-7B +GENTKG	**0.372**	**0.488**	0.563	0.217	**0.372**	0.443	**0.185**	**0.278**	**0.358**	0.715	0.767	0.786
	Llama2-7B - inductive	-	-	-	**0.234**	0.333	0.409	0.142	0.213	0.281	0.638	0.697	0.762

Table 2. Performance Comparison

Elizabeth Boschee, Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, James Starz, and Michael Ward. Icews coded event data. Harvard Dataverse, 12, 2015.
Kalev Leetaru and Philip A Schrodt. Gdelt: Global data on events, location, and tone, 1979–2012. In ISA annual convention, volume 2, pages 1–49. Citeseer, 2013.
Farzaneh Mahdisoltani, Joanna Biega, and Fabian M Suchanek. Yago3: A knowledge base from multilingual wikipedias. In CIDR, 2013.

SIEMENS | Munich Center for Machine Learning | LMU Ludwig-Maximilians-Universität München

Cross-domain Generalizability

- Single dataset evaluation. Exceeding performance on ICEWS14.
- Cross-checking across datasets:
  - **Slight performance drop** compared to the full training setting.
  - Even **higher performance than ICL** can be observed.

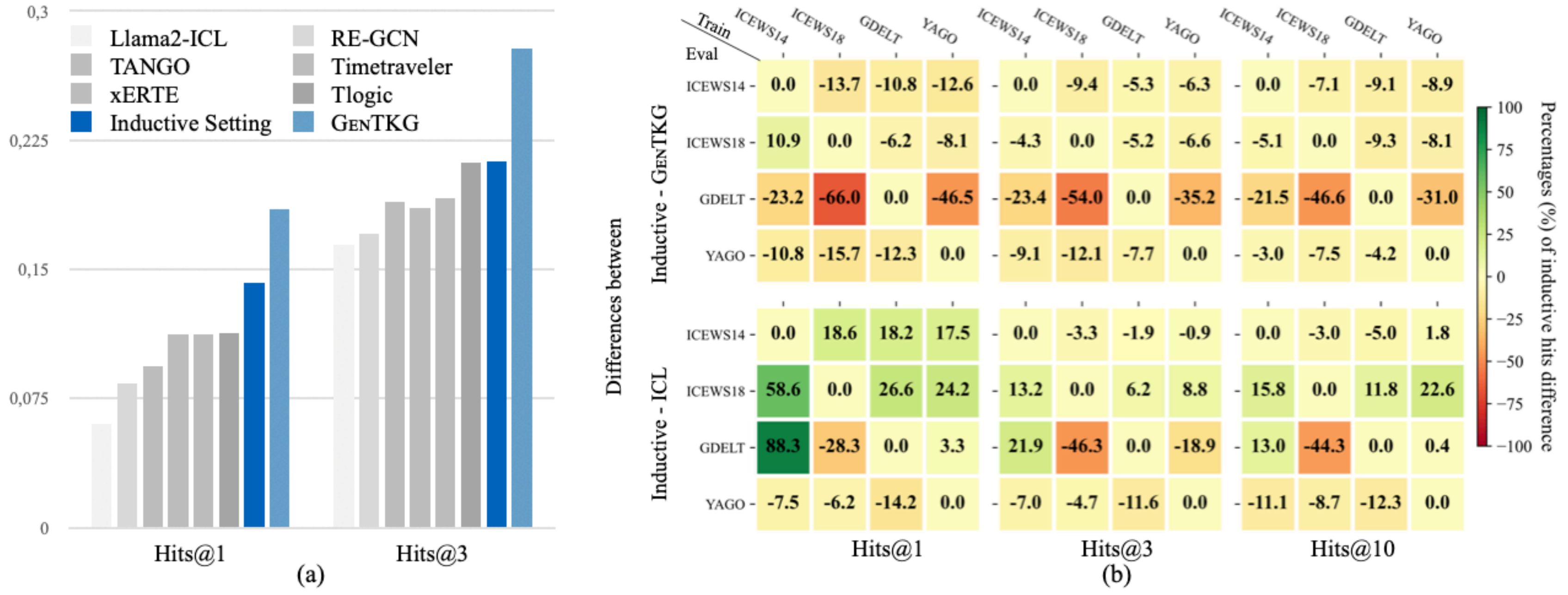


Figure 2. Investigating Inductive Setting comparing to conventional methods

Inductive Setting

In-domain Generalizability

- Trained on various partitions of training data of ICEWS14
- Exceeding performance on evaluation starting from the same timestamp.

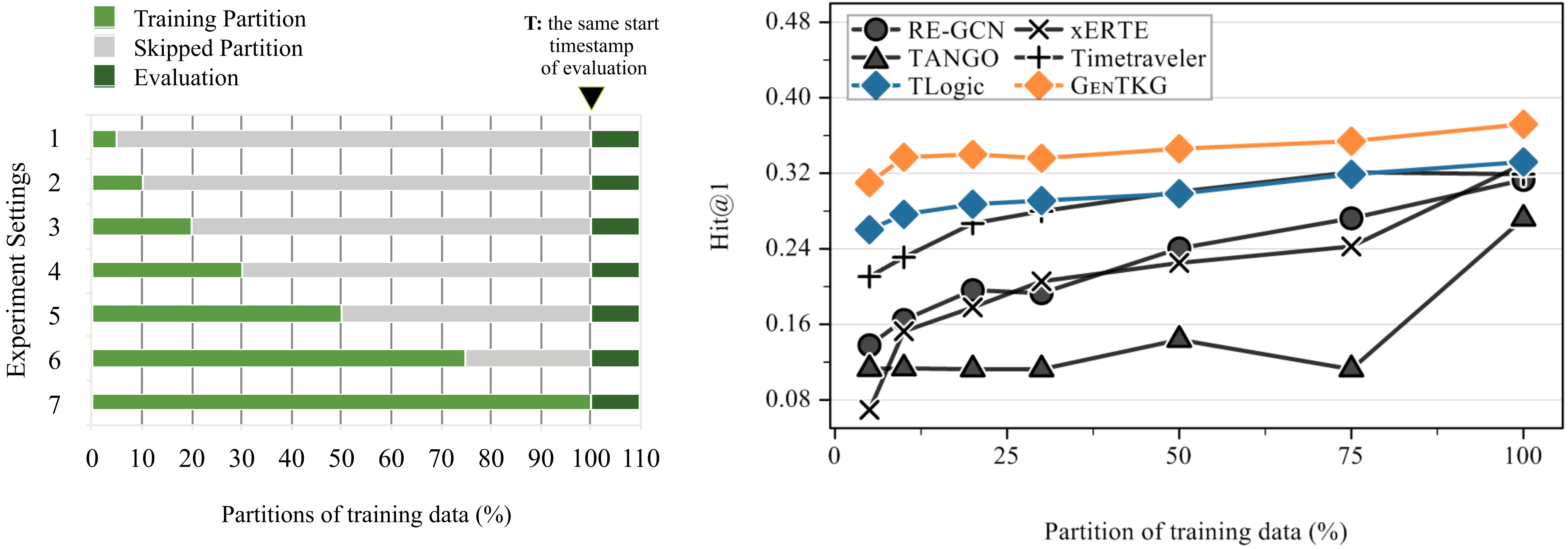
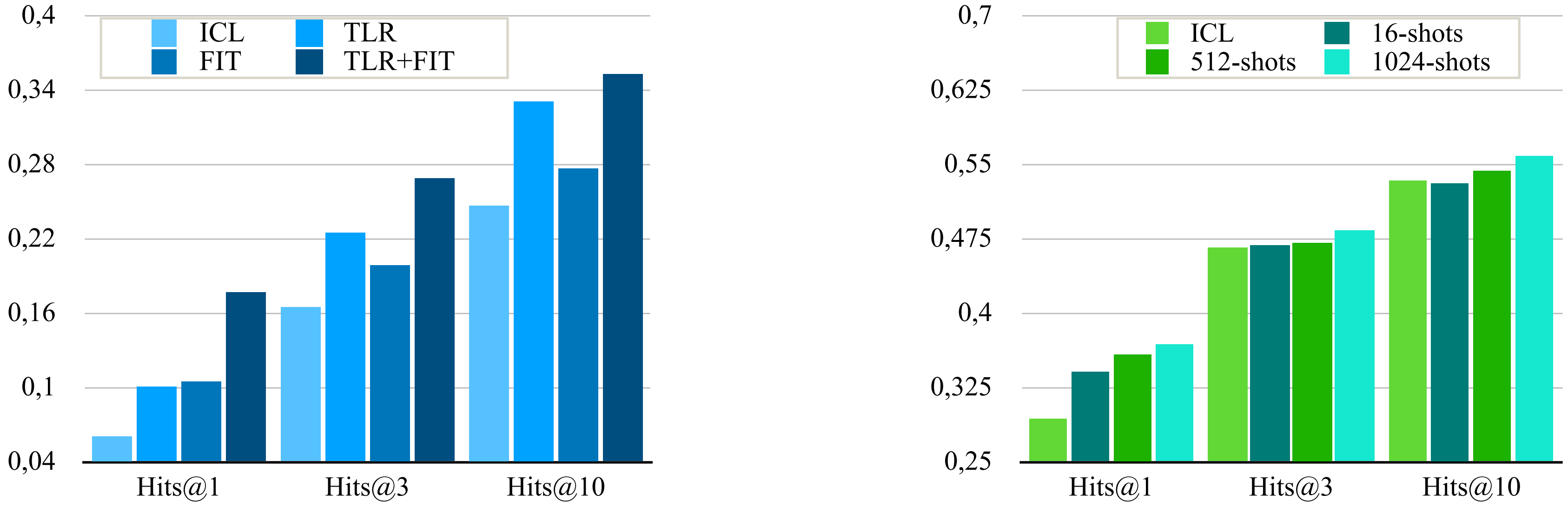


Figure 3. Exceeding performance of GenTKG in various partitions of training data.

Ablation Study

- **Both** TLR and FIT contribute to GenTKG.
- The **higher** the K, the **better** the performance.



(a) Ablation studies.

(b) Few-shot tuning.

Figure 4. Investigating TLR and FIT phase of GenTKG framework and the effect of K-shot training.

Contributions

- **Opening a frontier of generative forecasting on tKG.**
  - The first work that introduces instruction-tuned generative LLM to the tKG domain.
  - Retrieval augmented Generation Paradigm for tKG forecasting.
  - Regardless of the backbone LLM.
  - **Exceeding performance over conventional methods.**

- **Task reformulation from data learning to task alignment.**
  - Data-centric learning to task-centric LLM alignment.
  - Aligns LLMs to generative forecasting task on tKG.

- **Both Cross-domain and In-domain Generalizability without retraining.**
  - **Cross-domain generalizability**: one-time training on a single dataset with exceeding performance on multiple datasets without retraining.
  - **In-domain generalizability**: training on various partitions of training data of the same dataset with exceeding performance on the same evaluation.

- **Drastically low computation costs with exceeding performance.**
  - (0.0027%)**16**-shot tuning, comparable results to conventional methods.
  - (0.27%)**1024-**shots tuning, outperforming existing methods.

GenTKG: Generative Forecasting on Temporal Knowledge Graph

Ruotong Liao, Xu Jia, Yunpu Ma, Volker Tresp

ruotong.liao@outlook.com

16. Dec. 2023