Project Track:

Research

Team members:

Dayou Wu, dayouwu2@illinois.edu (Coordinator)
Kaiyue Zhang, kaiyuez2@illinois.edu
Bohan Liu, bohan3@illinois.edu

Research Question:

This research explores the feasibility of enhancing the temporal reasoning capabilities of large language models, such as GPT, which are known to possess inherent political biases and a limited ability to understand recently occurring events. Recognizing the high costs associated with training a new model from scratch to comprehend the latest happenings, this study proposes a novel approach. It involves leveraging the timely and dynamic nature of social media datasets along with other event data to construct a robust Temporal Knowledge Graph. This graph is then used to fine-tune the large language model, aiming to significantly improve its ability to reason about and understand temporal dynamics in the context of current events, thereby addressing its existing limitations in political neutrality and temporal awareness. The research problem will be defined as can we leverage the timeliness of social media datasets and other event data to construct an effective Temporal Knowledge Graph and subsequently fine-tune the LLM to enhance its temporal reasoning ability on contemporary events?

Significance:

Understanding the event reference in social media text data is beneficial to big data tasks, like user-based recommendation systems. Recent advances in large language models show the capacity of large transformer-based models in natural language understanding, but such models lack the capacity to understand time in numerical representations, and their perception of text data is subject to hallucination. Leveraging a robust event-time graph can help LLMs better resolve different events and accurately identify social trends.

Novelty:

The novelty lies in the specific use of dynamic, real-time data sources for TKG construction and applying this approach to improve political neutrality and temporal awareness in LLMs. Some existing works include Chain of History reasoning (Luo et. al., Xia et. al.), which emphasizes the instruction of LLMs to explore significant history chains and reason about possible answers for given queries, TKG-LM (Zhang et. al.), which emphasizes the need to exploit the semantic text information of entities and relations in TKGs and demonstrates how the integration of LLMs and graph neural networks can enhance temporal reasoning. These works have focused on structuring the reasoning process, integrating knowledge graphs with LLMs for enhanced prediction, and exploring the limitations and potentials of temporal reasoning in natural language processing. Our research could contribute to this field by offering a novel integration method and a focus on real-time data sources for temporal reasoning enhancement.

Approach:

To transform data from social media into a structured knowledge graph that can effectively train Large Language Models (LLMs), our approach encompasses several crucial steps. Firstly, we will preprocess the noisy data from social media to filter out spam and irrelevant noise, ensuring that the remaining data is of high quality and relevance. Next, we will mine this preprocessed data for valuable information, specifically focusing on detecting entity information and event mentions within the free text. This step is vital for extracting structured insights from unstructured data.

Following the data mining process, we will align the identified entities and events with those in the GDELT database. This alignment serves to enrich and complement the existing knowledge graph in GDELT with additional information and temporal insights not previously captured, thereby creating a more comprehensive and updated temporal knowledge graph.

Once these tasks are completed, we will conduct a preliminary evaluation to assess the effectiveness of our constructed knowledge graph in enhancing the temporal reasoning of LLMs. If time permits, and based on the outcomes of our initial evaluation, we will proceed to fine-tune the LLMs. This fine-tuning process will specifically aim to improve

the models' performance on temporal reasoning tasks, such as time-sensitive question answering, providing a clear demonstration of the enhanced capabilities of LLMs when augmented with our refined knowledge graph.

Evaluation:

For evaluation on Temporal Knowledge Graph, we will mainly focus on the Link Prediction jobs, time prediction, and entity prediction jobs. Mainly accuracy, precision, recall, and F-1 scores would be measured to judge our performance. For the temporal reasoning ability of fine-tuned LLMs, we will use the temporal reasoning benchmarks including MC-TACO, Time-Dial, and TimeQA to measure the performance on temporal based Question Answering tasks and time prediction jobs.

Timeline:

Week 1: Dataset, Preprocessing, Model Inference setup, investigate the information filtering model.
Week 2: Data mining, prompting models from week 1. Align social media text with time-annotated graph data.
Week 3: Fine-tune models from time-aligned data. Evaluation, Plot Tables, Run Baselines, Write up.

Task Division:

Kaiyue is responsible for setting up data preprocessing code, writing graph query interface, and fine-tuning LLM, and write-ups.
Bohan is responsible for setting up the model inference pipeline, temporal event postprocessing, and writeups. and running baselines, and write-ups.
Dayou is responsible for the model to filter useful information, writing temporal event extraction, and write-ups.

Reference:

Luo, R., Gu, T., Li, H., Li, J., Lin, Z., Li, J., & Yang, Y. (2024, February 14). *Chain of history: Learning and forecasting with LLMS for temporal knowledge graph completion*. arXiv.org. https://arxiv.org/abs/2401.06072

Xia, Y., Wang, D., Liu, Q., Wang, L., Wu, S., & Zhang, X. (2024, February 22). *Enhancing temporal knowledge graph forecasting with large language models via chain-of-history reasoning*. arXiv.org. https://arxiv.org/abs/2402.14382

Zhang, T., Zheng, T., Chen, Z., Li, L., Feng, Z., Zhang, D., & Song, M. (2023, October 13). *TKG-LM: Temporal Knowledge Graph Extrapolation Enhanced by Language Models*. OpenReview. https://openreview.net/forum?id=T0hhkuv8I0