



Thesis proposal

Topic: Investigating Temporal Bias in Pretrained Language Models

Supervisor: ZeinabSadat Taghavi

Examiner: Hinrich Schutze

Level: BSc / MSc

Summary: Introduction

LLMs like GPT, T5, and Llama have demonstrated exceptional capabilities across a range of natural language processing tasks. However, many real-world applications involve reasoning about time-dependent facts—information tied to specific timeframes, such as historical events, future predictions, or chronological ordering. Temporal reasoning, a critical component for such tasks, has been shown to expose significant limitations in these models.

A key challenge is temporal bias: PLMs often underperform when dealing with events from historical or less frequent time periods due to imbalanced token distributions in their pretraining corpora, which are dominated by contemporary data. This results in inaccurate temporal predictions, particularly for questions involving past or future events.

This thesis focuses on understanding and addressing temporal biases in LLMs. Specifically, it will:

1. Evaluate baseline performance of LLMs on temporal reasoning datasets like TEMPREASON and COMPLEXTMPQA. Finding the effect of variation of date formats on the model's ability to have temporal reasoning. (BSc)
2. Propose fine-tuning strategies using tasks such as temporal span extraction to improve performance across diverse time periods. (MSc)

By systematically investigating these issues, the thesis aims to enhance the temporal reasoning abilities of LLMs and contribute to the broader understanding of their limitations in handling time-sensitive information.

Requirements: Huggingface, being familiar with Transformer models.

References:

- Qingyu Tan, Hwee Tou Ng, and Lidong Bing (2023). *Towards Benchmarking and Improving the Temporal Reasoning Capability of Large Language Models*. DOI: 10.48550/ARXIV.2306.08952. URL: <https://arxiv.org/abs/2306.08952>
- Raphael Gruber et al. (2024). *ComplexTempQA: A Large-Scale Dataset for Complex Temporal Question Answering*. DOI: 10.48550/ARXIV.2406.04866. URL: <https://arxiv.org/abs/2406.04866>