Time-Aware Retrieval-Augmented Generation

[Group 11]: Gauri Patki, Sonith Bingi, Prasad Jawale, Manoj Arulmurugan

Problem Description

- Traditional RAG models retrieve data only based on semantic relevance and that can lead to irrelevant/outdated information
- The models fail for time-sensitive tasks
- Our Approach: Consider both semantic and temporal relevance to fetch documents

Significance & Research Value

- Help develop systems that distinguish between outdated and current knowledge
- Enable any knowledge based system (Eg: Legal, Finance QA) to give up-to-date answers

Example:
Which countries hosted the Olympics between 2000 and 2016 but not after 2012?

Literature Survey & Comparison with Proposed Method

Proposed Methods or Explorations

- **Variant A (Time-aware retriever):** Keep MRAG's MC/TC split, but add a small time vector to query/doc embeddings and apply a first-stage time filter
- **Why A:** Top-k is on-topic + time-consistent from the start; optional finetune with temporal hard negatives for extra recall.
- **Variant B:** Offline tag chunks with time intervals → pre-filter by TC (as-of/before/after/between), then rank with a time-anchored query (avg. of "In Jan 2014, In Feb 2014" variants).
- **Why B:** Blocks post-cutoff leakage and steers results toward the cutoff: training-free, quick.

Technical Challenges

- **Date ambiguity:** facts mention multiple years; naïve "find a year" rules pick the wrong one.
- **Publication vs event time:** page update time ≠ fact validity time.
- **Sparse/absent dates:** many chunks have no explicit dates.

Experiment Designs and Feasibility Analysis

Datasets and Evaluation Metrics

- TempRAGEval (TimeQA + SituatedQA: Wikipedia 2021 Dump)
- TORQUE Dataset
- Evaluation Metric - Exact Match(EM) & Answer Recall

Baseline Methods

- Standard RAG model without temporal features and compare it with a temporal RAG
- MRAG Benchmarks on the TempRAGEval dataset

Computing Estimation

- 1-2 NVidia A40 48GB GPUs
- Time extraction + text retrieval using LLaMA: ~30 Minutes

Model Checkpoints and Codebase

- LLaMA 2-7B model

https://arxiv.org/pdf/2412.15540
https://arxiv.org/pdf/2005.00242