

# Weekly Report

Wang Tianyi

October 29, 2024



# Outline

- 1 Experiments
  - Main experiment
  - Accuracy Table

# Outline

- 1 Experiments
  - Main experiment
  - Accuracy Table

# Experiments

- 1 We conducted experiments using a benchmark dataset with 775 test questions, and the queries were generated based on different prompts. The experiments varied by the form of the query and the matching methods used.
- 2 **Experiments without weighting for abstract and mechanism:**
  - Long Sentence Query + Cosine Similarity Matching: Achieved an accuracy of 50%.
  - Long Sentence Query + CLS Matching: Achieved an accuracy of 48%.
  - Multi-Token Query + CLS Matching: Achieved an accuracy of 48%.
  - Multi-Token Query + Cosine Similarity Matching: Achieved an accuracy of 64%.
- 3 **Experiments with abstract weight 0.2 and mechanism weight 0.8:**
  - Multi-Token Query + CLS Matching (with abstract as input for check and summarize stages): Achieved an accuracy of 53%.
  - Long Sentence Query + CLS Matching (with abstract as input for check and summarize stages): Achieved an accuracy of 46.46%.
  - Long Sentence Query + CLS Matching (with both abstract and mechanism as input for check and summarize stages): Achieved an accuracy of 51.52%.

## Accuracy Comparison for Different Query Forms and Matching Methods

Experiment Setting	Matching Method (Accuracy %)	
	Cosine Similarity	CLS Matching
<b>No Weighting for Abstract and Mechanism</b>		
Long Sentence Query	50.00	48.00
Multi-Token Query	<b>64.00</b>	48.00
<b>With Abstract Weight 0.2, Mechanism Weight 0.8</b>		
Multi-Token Query (Abstract as Input)	-	53.00
Long Sentence Query (Abstract as Input)	-	46.46
Long Sentence Query (Abstract+Mechanism)	-	51.52