

CS 421 Reproducibility Study: *EchoPrompt: Instructing the Model to Rephrase Queries for Improved In-context Learning*

Davide Salonic (UIN: 661296414)

University of Illinois at Chicago

- Paper: *EchoPrompt: Instructing the Model to Rephrase Queries for Improved In-context Learning*.
- Goal: **Enhance in-context learning for LLMs by using rephrased and repeated queries.**
- Key claim: Improved accuracy in tasks such as numerical reasoning and reading comprehension.

Scope of Reproducibility

- Focused on validating the following claim: EchoPrompt improves both zero-shot and few-shot accuracy for different reasoning tasks.
- Reproduced on a subset of dataset and models

- **Model:** GPT-3.5-turbo accessed via OpenAI API.
- **Datasets:**
 - MultiArith: Multi-step numerical reasoning problems.
 - DROP_break: Discrete reasoning tasks from paragraphs.
- **Process:**
 - Generated rephrased queries with task-agnostic prompts.
 - Combined original and rephrased queries for final reasoning.
- **Computational Setup:**
 - Cloud-based API
 - 10 hours runtime
- **Code:** EchoPrompt GitHub.

Example of Prompt

Q: Kelly has 5 quarters and 2 dimes. If she buys a can of pop for 55 cents, how many cents will she have left?

Zero-Shot CoT:

A: Let's think step by step. Kelly has 5 quarters and 2 dimes.
So, she has 5 quarters and 2 dimes left.
 $5 \text{ quarters} = 5 \times 25 \text{ cents} = 125 \text{ cents}$
and $2 \text{ dimes} = 2 \times 10 \text{ cents} = 20 \text{ cents}$.
So, she has $125 + 20 = 145 \text{ cents}$ left.
Therefore, the answer is 145 cents. ×

Zero-Shot CoT + EchoPrompt:

A: Let's repeat the question and also think step by step. Kelly has 5 quarters and 2 dimes. She buys a can of pop for 55 cents. How many cents will she have left?

Let's start with the first step. Kelly has 5 quarters and 2 dimes. We can represent this as:

$5 \text{ quarters} = 5 \times 25 \text{ cents} = 125 \text{ cents}$.

$2 \text{ dimes} = 2 \times 10 \text{ cents} = 20 \text{ cents}$.

So, Kelly has $125 + 20 = 145 \text{ cents}$.

She buys a can of pop for 55 cents. We can represent this as: 55 cents.

So, Kelly has $145 - 55 = 90 \text{ cents}$ left.

Therefore, the answer is 90 cents. ✓ 🔍 ↺

Reproduced results confirm original claims:

| Dataset | Data source | Zero shot Standard | Zero shot CoT | Few shots Standard | Few shots CoT |
|------------|-----------------------|--------------------|------------------|--------------------|-----------------|
| MultiArith | Original Paper | 31.0/48.5(+17.5) | 76.0/78.7(+2.7) | 44.0/53.8(+9.8) | 96.1/97.8(+1.7) |
| MultiArith | Reproducibility Study | 31.2/47.5(+16.3) | 75.9/78.8(+2.9) | 44.0/53.5(+9.5) | 96.1/97.8(+1.7) |
| DROP_break | Original Paper | 43.7/55.8(+12.1) | 38.2/51.2(+13.0) | 55.5/63.1(+7.6) | 65.3/69.6(+4.3) |
| DROP_break | Reproducibility Study | 44.1/55.4(+11.3) | 40.0/51.3(+11.3) | 55.6/63.4(+7.8) | 65.2/69.7(+4.5) |

Accuracy always improves when using EchoPrompt

Challenges and Future Work

- **Challenges:**

- Deprecated models (e.g., code-davinci-002) limited reproducibility.
- Budgetary constraints on API usage.

- **Observation:** Model updates (weights) may lead to slight deviations from original results.

- **Future Work:**

- Test additional datasets and models.
- Investigate systematic prompt engineering for further optimization.

Thank You!
Questions?