FGCU

U.A. WHITAKER COLLEGE OF ENGINEERING

Dendritic: A Human-Centered Artificial Intelligence and Data Science Institute

Introduction to Prompt Engineering

John Peller

May 12, 2025

The sky is…

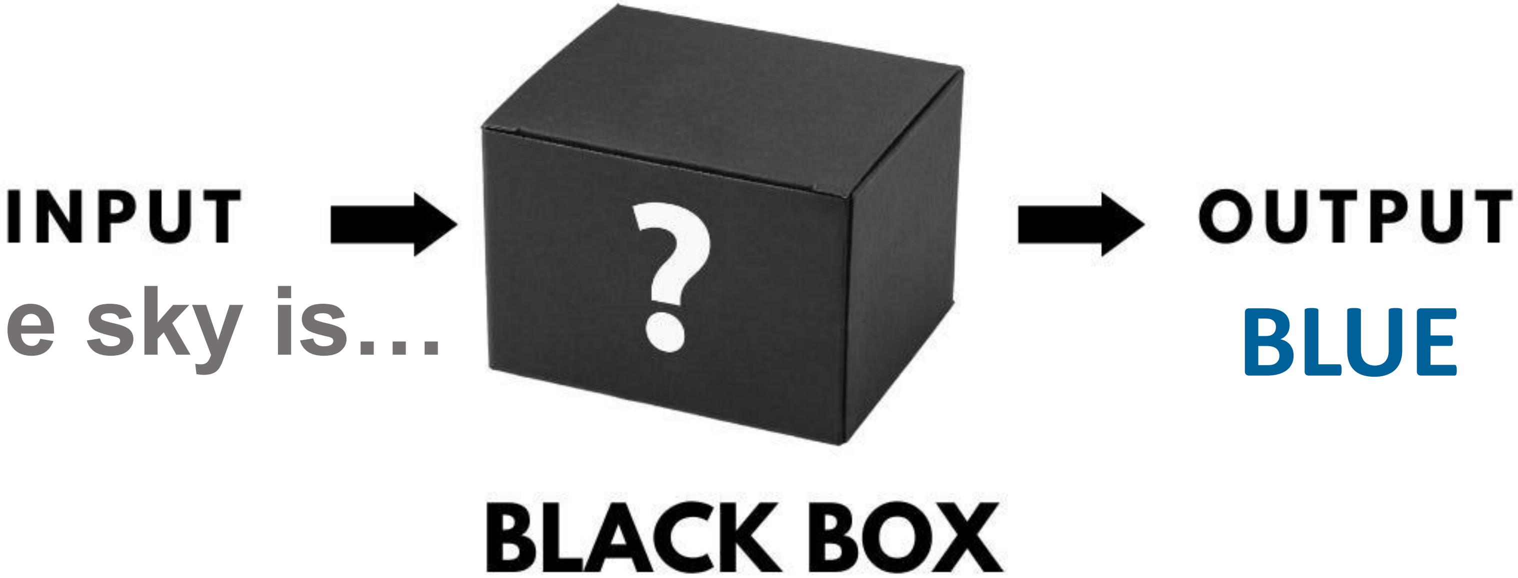
The sky is…

BLUE

The sky is…

BLUE	the limit
falling	up
overcast	clear today
full of stars	the daily bread of the eyes

How an LLM ‘thinks’

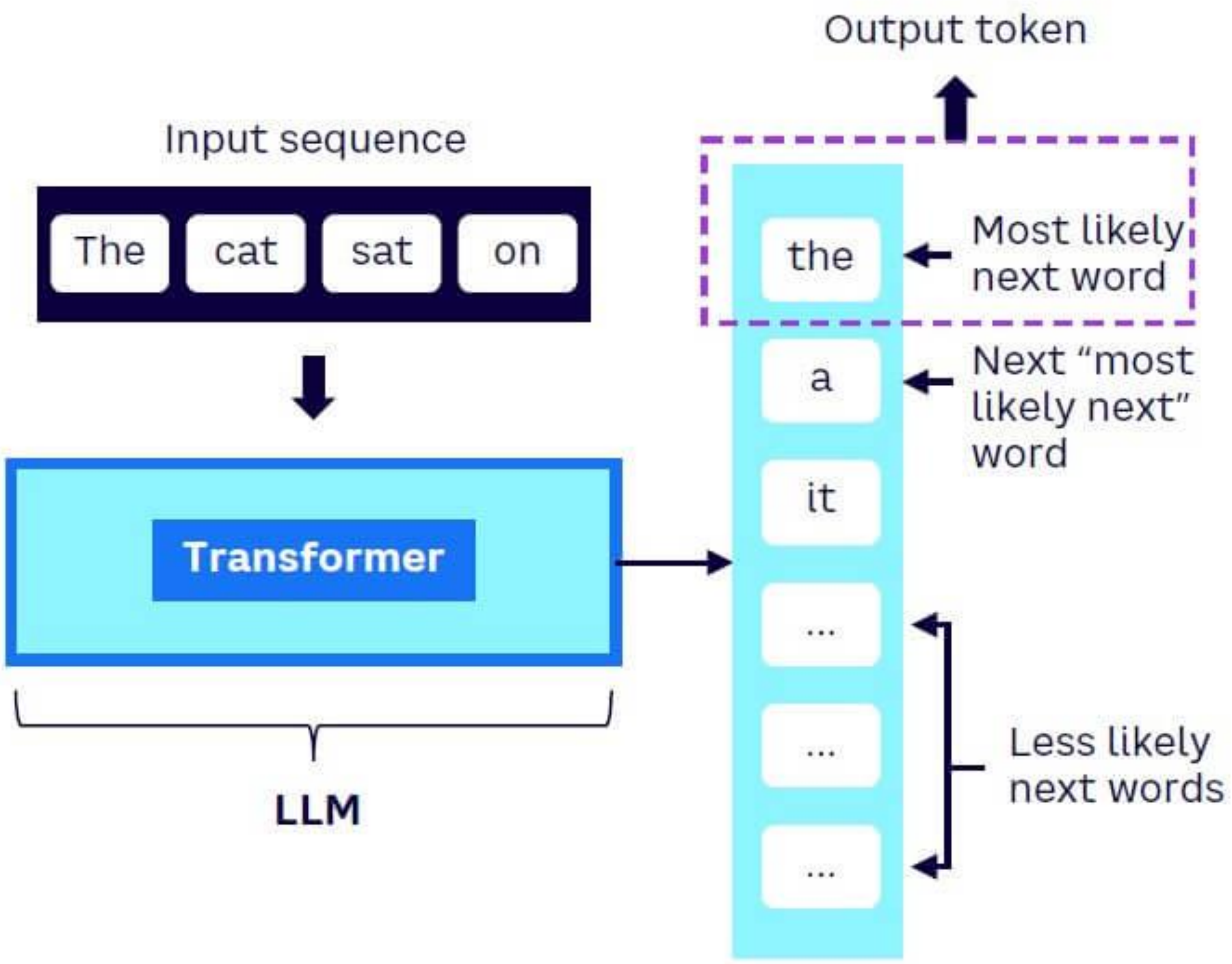


The sky is…

BLUE

FGCU

How an LLM 'thinks'



Source: "Introduction to Large Language Models." Google Cloud, 2023

How an LLM ‘thinks’

Prompt				Generated next word
Four score and seven	→	GPT	→	*years*
Four score and seven *years*	→	GPT	→	*ago*
Four score and seven *years ago*	→	GPT	→	*our*
Four score and seven *years ago our*	→	GPT	→	*fathers*
Four score and seven *years ago our fathers …. Created equal*	→	GPT	→	*<E>*

https://www.columbusconsulting.com/insights/simply-defining-ai-artificial-intelligence/

FGCU

How an LLM 'thinks'

Prompt

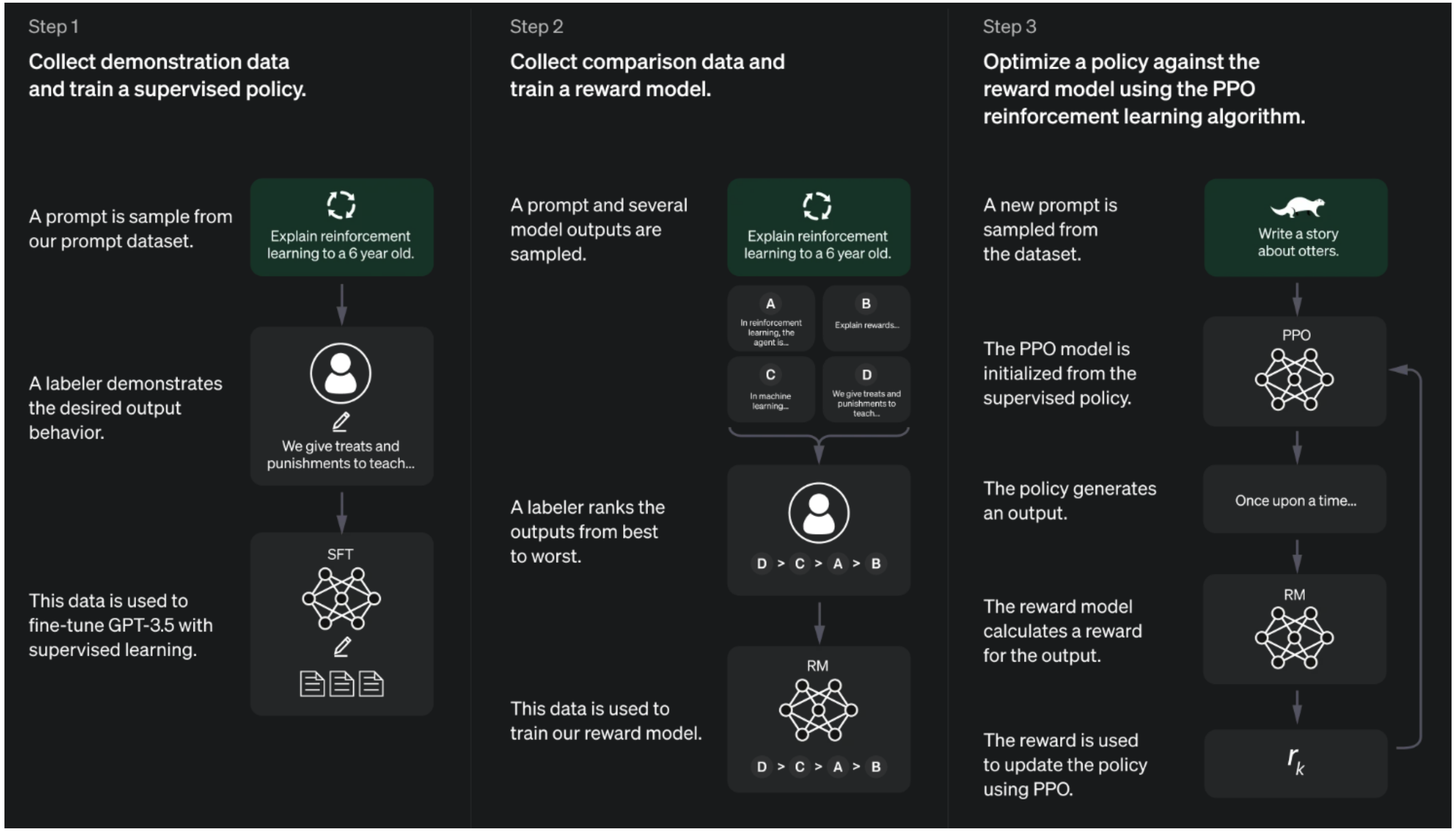
Four score and seven → **GPT** →

Possible final output

Four score and seven years ago our fathers brought forth, upon this continent, a new nation, conceived in liberty, and dedicated to the proposition that all men are created equal.

https://www.columbusconsulting.com/insights/simply-defining-ai-artificial-intelligence/

How an LLM ‘thinks’



https://openai.com/index/chatgpt/

How do we align our prompt with the model?

How do we align our prompt with the model?

- In a word, Context

How do we align our prompt with the model?

- In a word, Context
- In depth, Prompt Engineering
  - **Instruction** - state the task you want the model to perform
  - **Context** - background information, guidelines, or constraints
  - **Input Data** - specific text/data the model should work on
  - **Format** - how you want the answer returned: structure, style, etc.

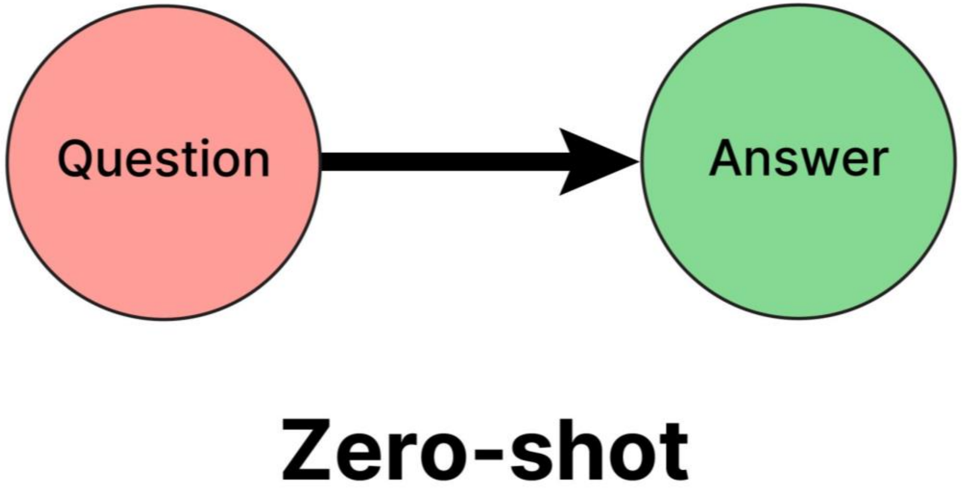
FGCU

How do we align our prompt with the model?

- **Instruction**
  - Analyze customer feedback and classify sentiment.
- **Context**
  - You are a customer-experience analyst at ACME Corp. Provide structured insight into user comments to guide product improvements. Use neutral, professional language.
- **Input Data**
  - 1) “My rocket skates blew up!”
  - 2) “Anvil arrived by airmail, directly on target.”
  - 3) “Portable hole only worked for roadrunner.”
- **Format**
  - Return a JSON array where each element consists of:
    - `id`: the comment number
    - `sentiment`: “positive”, “negative”, or “neutral”
    - `key_phrases`: exactly one concise phrase capturing main point

Zero-Shot Prompting

- Directly instruct the model on a task
- Establishes baseline of the model's innate utility and helps identify where examples or further guidance are needed

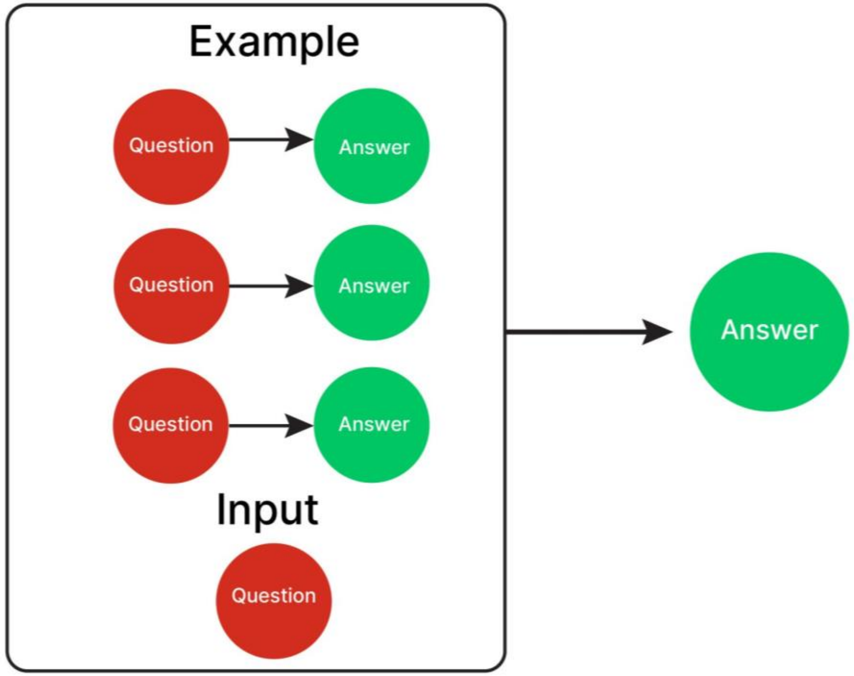


https://www.analyticsvidhya.com/blog/2024/06/zero-shot-prompting/

FGCU

Multi-Shot Prompting

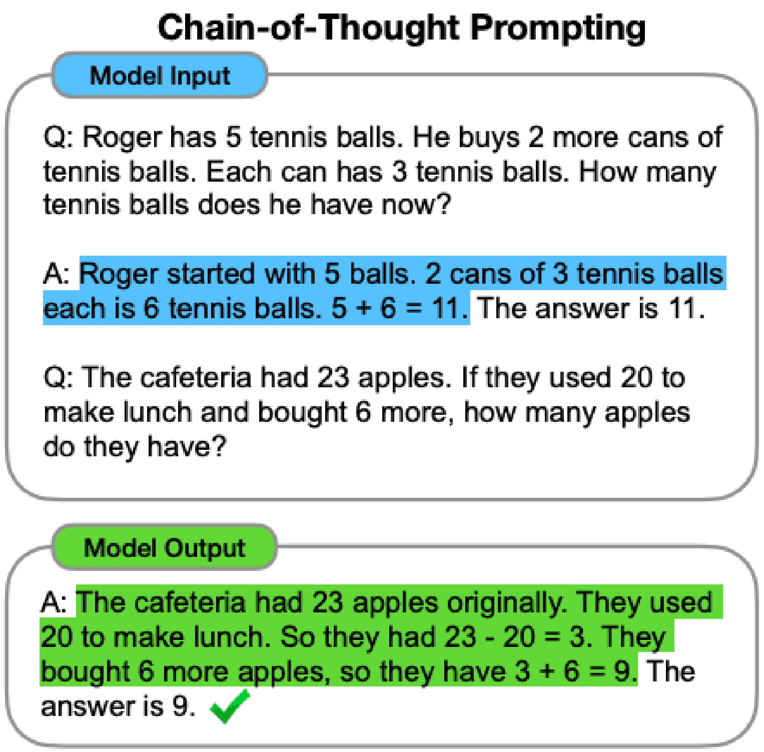
- Provide a small set of input–output pairs within the prompt to demonstrate the desired format
- Gives the model pattern induction or disambiguation



https://www.analyticsvidhya.com/blog/2024/07/few-shot-prompting/

Chain-of-Thought Prompting

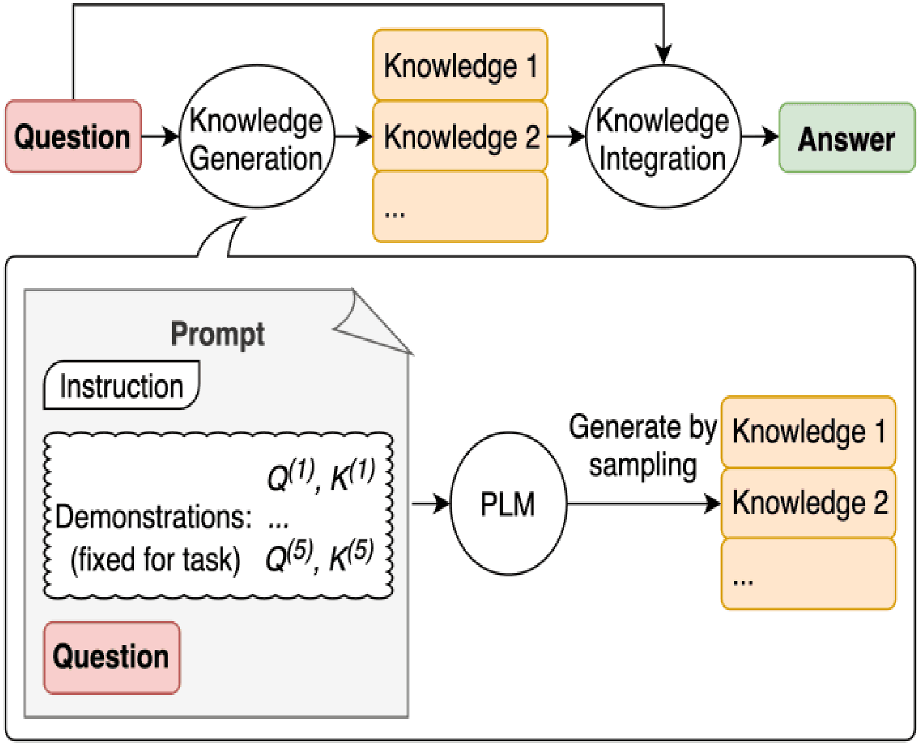
- Articulate intermediate reasoning steps before arriving at an answer
- Stepwise improves accuracy in complex arithmetic, logic puzzles, and reasoning challenges



https://www.promptingguide.ai/techniques/cot

Generate Knowledge Prompting

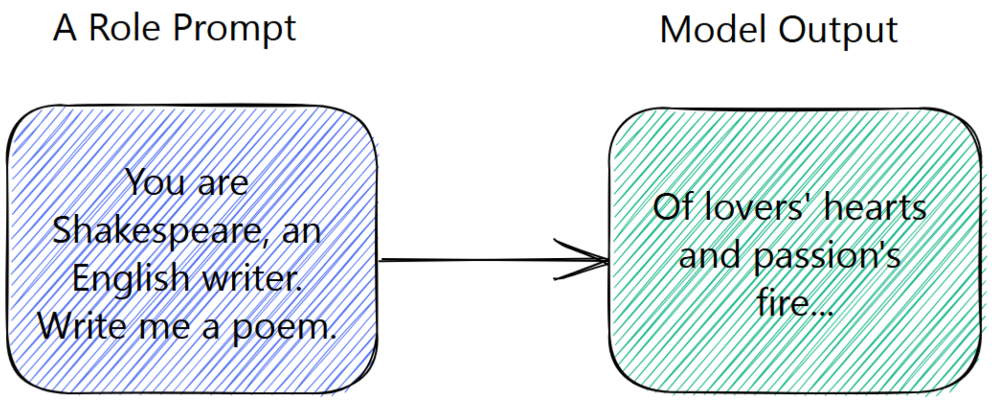
- First prompt the model to list relevant facts or background knowledge, then uses that foundation to answer a question
- Anchors model answer in explicitly generated content



https://arxiv.org/abs/2110.08387

Role Prompting

- Instruct the model to adopt a specific persona or viewpoint before handling the main task
- Narrows generative scope and aligns responses with domain conventions



https://learnprompting.org/docs/basics/roles

Putting it Together

Multi-Shot

- Columns – example1 – example2:
  - Id – 1 – 2
  - SepalLengthCm – 5.1 – 4.9
  - SepalWidthCm – 3.5 – 3
  - PetalLengthCm – 1.4 – 1.4
  - PetalWidthCm – 0.2 – 0.2
  - Species – Iris-setosa – Iris-setosa

Chain-of-Thought

- Provide stepwise reasoning:
  - Compare how feature distributions differs across species.
  - Identify which single feature yields cleanest separation.
  - Consider combinations of features for improved accuracy

Generate Knowledge

- List three key statistical features of sepal and petal measurements that can be used to clearly distinguish the three Iris species.

Role Prompting

- “You are acting as an experienced botanical data scientist, evaluating the Iris dataset for a non-technical audience”

Final Prompt

- Propose a simple rule (or set of rules) to classify an unlabeled Iris sample into one of the three species and explain why your rule works.

FGCU

Prompt Experiment Homework

You support Wildfire Risk Analysis at Montesinho Natural Park in Portugal. The superintendent needs an assessment for resource planning.

Task: Predict the fire risk level for this observation based on historical correlation with area (burned area) in the forestfires dataset:

- | X | Y | month | day | FFMC | DMC | DC | ISI | temp | RH | wind | rain | area |
- | 7 | 5 | mar | fri | 86.2 | 26.2 | 94.3 | 5.1 | 8.2 | 51 | 6.7 | 0 | 0 |

Step 1 - Chain-of-Thought (show your work):

- 1. Compare and discuss FFMC (Fine Fuel Moisture Code): <80 = low, 80–90 = moderate, >90 = high.
- 2. Compare and discuss DMC (Duff Moisture Code): <15 = low, 15–30 = moderate, >30 = high.
- 3. Compare and discuss DC (Drought Code): <60 = low, 60–120 = moderate, >120 = high.
- 4. Compare and discuss ISI (Initial Spread Index): <3 = low, 3–6 = moderate, >6 = high.
- 5. Summarize weather factors: temp, RH, Wind, Rain.
- 6. Integrate all factors (assign appropriate weighting) to infer overall risk.

Step 2 - Analysis: Provide a paragraph discussing primary contributing factors, identifying calculated historical correlations with burned area.

Step 3 - Prediction: Based on all of the above, provide your final predicted fire risk level of the provided observation: low/medium/high.

FGCU

Advice

- Vary and iterate
- Be aware of hallucinations
- Reduce verbosity
- Be polite*

*maybe

Should We Respect LLMs? A Cross-Lingual Study on the Influence of Prompt Politeness on LLM Performance

Ziqi Yin[1] **Hao Wang**[1] **Kaito Horio**[1] **Daisuke Kawahara**[1,2,3] **Satoshi Sekine**[2,3]
[1]Waseda University [2]RIKEN AIP [3]NII LLMC
{yinziqi2001@toki.,conan1024hao@akane.,kakakakakakaito@akane.,dkw@}waseda.jp
satoshi.sekine@riken.jp

Abstract

We investigate the impact of politeness levels in prompts on the performance of large language models (LLMs). Polite language in human communications often garners more compliance and effectiveness, while rudeness can cause aversion, impacting response quality. We consider that LLMs mirror human communication traits, suggesting they align with human cultural norms. We assess the impact of politeness in prompts on LLMs across English, Chinese, and Japanese tasks. We observed that impolite prompts often result in poor performance, but overly polite language does not guarantee better outcomes. The best politeness level is different according to the language. This phenomenon suggests that LLMs not only reflect human behavior but are also influenced by language, particularly in different cultural contexts. Our findings highlight the need to factor in politeness for cross-cultural natural language processing and LLM usage.

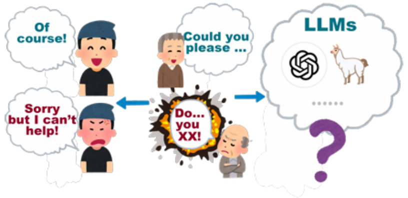


Figure 1: Illustration of our motivation.

in our language and behavior. However, politeness and respect may have different definitions and manifestations in different cultures and languages. For example, the expression and degree of respect in English, Chinese, and Japanese may differ significantly. This difference may make the performance of LLMs vary with language on the same politeness level.

We hypothesize that impolite prompts may lead to a deterioration in model performance, including

https://arxiv.org/abs/2402.14531

Works Cited

- Brown, Tom B., et al. “Language Models Are Few-Shot Learners.” arXiv Preprint arXiv:2005.14165v4, 2020. https://doi.org/10.48550/arXiv.2005.14165.
- Kong, Aobo, et al. “Better Zero-Shot Reasoning with Role-Play Prompting.” arXiv Preprint arXiv:2308.07702v2, 15 Aug. 2023, https://doi.org/10.48550/arXiv.2308.07702.
- Liu, Jiacheng, et al. “Generated Knowledge Prompting for Commonsense Reasoning.” arXiv Preprint arXiv:2110.08387v3, 2022. https://doi.org/10.48550/arXiv.2110.08387.
- Saravia, Elvis. “Prompt Engineering Guide.” GitHub, Dec. 2022, https://github.com/dair-ai/Prompt-Engineering-Guide.
- Wei, Jason, et al. “Chain-of-Thought Prompting Elicits Reasoning in Large Language Models.” arXiv Preprint arXiv:2201.11903v6, 2023. https://doi.org/10.48550/arXiv.2201.11903.
- Yin, Ziqi, et al. “Should We Respect LLMs? A Cross-Lingual Study on the Influence of Prompt Politeness on LLM Performance.” arXiv Preprint arXiv:2402.14531v2, 2024. https://doi.org/10.48550/arXiv.2402.14531.

FGCU

U.A. WHITAKER COLLEGE OF ENGINEERING

Dendritic: A Human-Centered Artificial Intelligence and Data Science Institute

Thank You!