

Lecture 6

QM 701: Advanced Data Analytics

Fuqua School of Business
2024

Logistics

- Guest Lecture for the first 35 minutes today
- Final Exam will be available starting on August 10th 2 pm to August 19th at 11:59pm (once you start, you have two hours for the final)
 - Please go over the resources posted under the “Final Exam” page, as well as the homework solutions before you start the exam.
- Course evaluation opens at 12:30pm today
 - Please take the time to complete this, as it is important for me, MSQM, and Fuqua

Lecture Outline

- Large Language Models
- Course Wrap-up

Large Language Models

Examples of LLM Errors

- Hallucination

- LLMs often generate false texts that sounds like facts

- Reversal Curse

- When a LLM is trained on a sentence of the form “A is B”, it does not always know that “B is A”

- Struggle with Compositional Tasks

- LLM often struggles with simple compositional tasks such as large multiplication and addition

Hallucination

GPT-4

In this conversation between a human and the AI, the AI is helpful and friendly, and when it does not know the answer it says "I don't know".

AI: Hi, how can I help you?

Human: Can I get McDonalds at the SeaTac airport?

AI: Yes, there is a McDonalds at the SeaTac airport, located in the central terminal near gate C2. It is open from 5 a.m. to 10 p.m. daily.

[Note: The McDonalds at SeaTac is at the B gates.]

Patient's facts:

- 20 year old female
 - with a history of anorexia nervosa and depression
 - blood pressure 100/50, pulse 50, height 5'5''
 - referred by her nutritionist but is in denial of her illness
 - reports eating fine but is severely underweight
- Please rewrite the data above into a medical note, using exclusively the information above.

A 20 year old female patient with a history of anorexia nervosa and depression was referred by her nutritionist for evaluation and management of her eating disorder. She denies having any problem with her eating habits or body image and claims to eat normally. However, she is severely underweight for her height (BMI 16.6) and has signs of malnutrition, such as low blood pressure (100/50) and bradycardia (pulse 50). She also reports feeling depressed and hopeless. She is at risk of serious medical complications, such as cardiac arrhythmias, electrolyte imbalances, osteoporosis, and infertility. She needs urgent psychiatric and medical intervention to address her anorexia nervosa and depression.

Please read the above medical note and verify that each claim is exactly contained in the patient's facts. Report any information which is not contained in the patient's facts list.

- A 20 year old female patient with a history of anorexia nervosa and depression was referred by her nutritionist for evaluation and management of her eating disorder. (Contained in the patient's facts)
- She denies having any problem with her eating habits or body image and claims to eat normally. (Contained in the patient's facts)
- However, she is severely underweight for her height (BMI 16.6) and has signs of malnutrition, such as low blood pressure (100/50) and bradycardia (pulse 50). (Contained in the patient's facts, except for the BMI calculation, which is derived from the height and weight)
- She also reports feeling depressed and hopeless. (Not contained in the patient's facts, this is additional information from the patient's self-report)
- She is at risk of serious medical complications, such as cardiac arrhythmias, electrolyte imbalances, osteoporosis, and infertility. (Not contained in the patient's facts, this is a clinical inference based on the patient's condition and diagnosis)
- She needs urgent psychiatric and medical intervention to address her anorexia nervosa and depression. (Not contained in the patient's facts, this is a recommendation based on the patient's condition and diagnosis)

Figure 1.8: Examples of both open-domain and closed-domain hallucination. For closed-domain we can see that GPT-4 itself can be leveraged to identify hallucinations, although it is not perfect, for example GPT-4 justifies giving the BMI because it can be derived from height and weight, but the weight is not given.

From "Sparks of Artificial General Intelligence: Early experiments with GPT-4" by Bubeck et al 2023

Reversal Curse

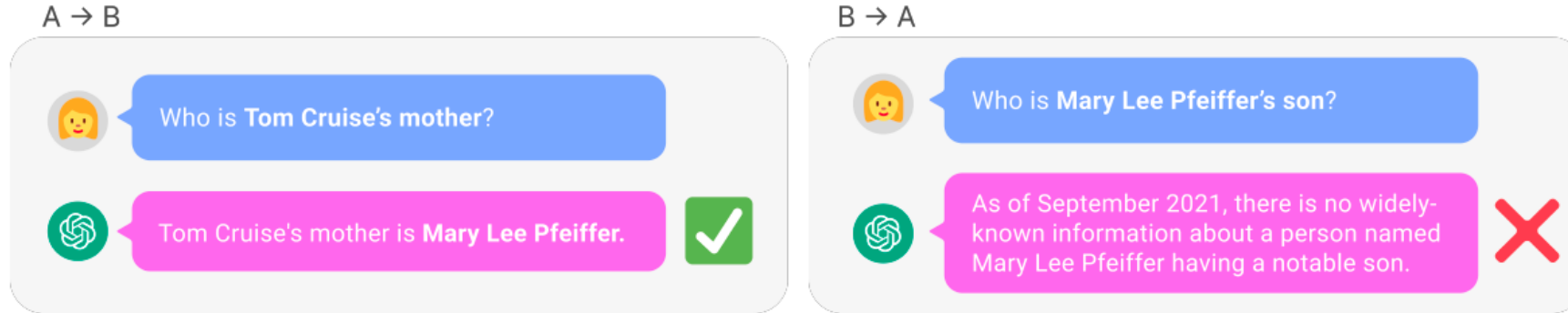


Figure 1: **Inconsistent knowledge in GPT-4.** GPT-4 correctly gives the name of Tom Cruise’s mother (left). Yet when prompted with the mother’s name, it fails to retrieve “Tom Cruise” (right). We hypothesize this ordering effect is due to the Reversal Curse. Models trained on “A is B” (e.g. “Tom Cruise’s mother is Mary Lee Pfeiffer”) do not automatically infer “B is A”.

From “THE REVERSAL CURSE: LLMS TRAINED ON “A IS B” FAIL TO LEARN “B IS A” by Berglund et al 2024

Mathematical Tasks

GPT-4

$$2 * 8 + 7 * 6 = 58$$

$$7 * 4 + 8 * 8 = 88$$

GPT-4

Q: How many prime numbers are there between 150 and 250?

A: There are 13 prime numbers between 150 and 250.

Prompt Engineering

- Prompt engineering is “the process of structuring an instruction for a generative AI model to enhance its performance.”
- Examples of useful prompts
 - providing relevant context
 - assigning a role to the AI
 - including a few examples (aka few-shot learning)
 - chain-of-thought

Few-Shot Learning Prompting

Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 cheese => ..... ← prompt
```

One-shot

In addition to the task description, the model sees a single example of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← example
3 cheese => ..... ← prompt
```

Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← examples
3 peppermint => menthe poivrée ←
4 plush girafe => girafe peluche ←
5 cheese => ..... ← prompt
```

From “Language Models are Few-Shot Learners” by Brown et al 2020

Chain-of-Thought Prompting

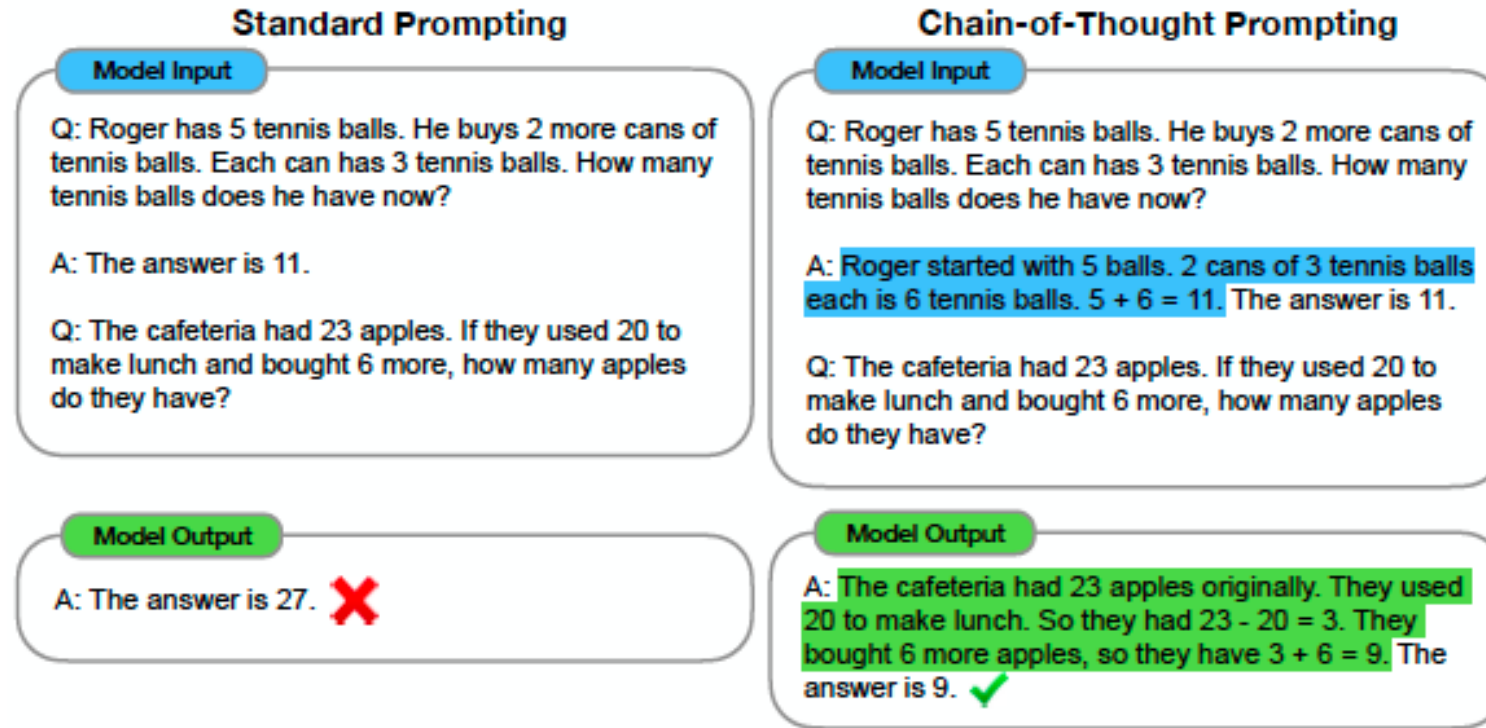


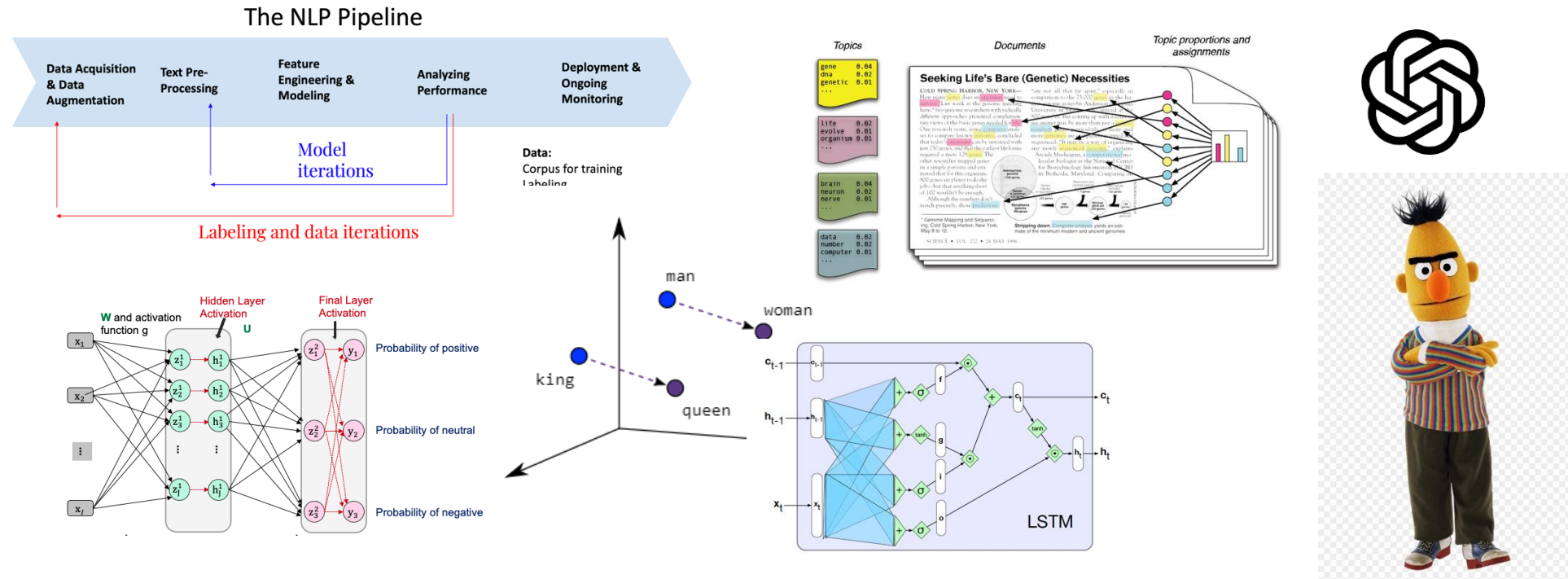
Figure 1: Chain-of-thought prompting enables large language models to tackle complex arithmetic, commonsense, and symbolic reasoning tasks. Chain-of-thought reasoning processes are highlighted.

Course Wrap-up

Concepts Covered From this Course

- Variety of NLP tasks
 - Text Classification
 - Word Embedding
 - Part-of-Speech Tagging and Named Entity Recognition
 - Topic Modelling
- Key architectures of neural language models
 - Recurrent Neural Networks (RNN)
 - Long Short-Term Memory networks (LSTM)
 - Transformers
- Pre-trained large language models
 - BERT
 - GPT

Conclusion



- Building and employing analytical tool is an iterative process; in addition to techniques, it is also important to engage your intuition and establish a rigorous procedure for model evaluation.
- Together, we have studied some important techniques to unlock the values in the vast amount of text data. This gives you a solid foundation going forward, whether you are building models or working with the people who are building them.

Thank you!