

EXPERIENCES TEACHING WITH / ABOUT LLMS



DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY

LLM, GPT, CREATIVITY
8 AUGUST 2023

ARTHUR HJORTH
ASSISTANT PROFESSOR



BRIEF OVERVIEW OF ACTIVITIES

- Learning about LLMs: What do students *need* to know?
- Three different kinds of activities that I have built/run:
 - Design your own Chatbot
 - Coding Assistance
 - Prompt Engineering
- Summarizing didactic thoughts

0 LLMS

Students should have *some* understanding of how LLMs work. What this “some” is, is obviously still contested.

My best evaluation at this point is that they should understand how “knowledge” of:

1. language, syntax, semantics
2. frequencies of word/phrase use
3. context

Is embedded in LLMs and that this is how they make predictions, and

That keeping this in mind when using LLMs makes for better, more desirable outputs

0 LLMS

Students should have *some* understanding of how LLMs work. What this “some” is, is obviously still contested.

My best evaluation at this point is that they should understand how “knowledge” of:

1. language, syntax, semantics
2. frequencies of word/phrase use
3. context

Is embedded in LLMs and that this is how they make predictions, and

That keeping this in mind when using LLMs makes for better, more desirable outputs

Let's guess words!

The red ____ rolled down the hill

0 LLMS

Students should have *some* understanding of how LLMs work. What this “some” is, is obviously still contested.

My best evaluation at this point is that they should understand how “knowledge” of:

1. language, syntax, semantics
2. frequencies of word/phrase use
3. context

Is embedded in LLMs and that this is how they make predictions, and

That keeping this in mind when using LLMs makes for better, more desirable outputs

Popular choices:
Ball, Car

What do we infer from context?

It *must* be a noun, referencing a rollable object

It is *probably* something that is typical of this situation.

What could ___ be, but is probably not?

0 LLMS

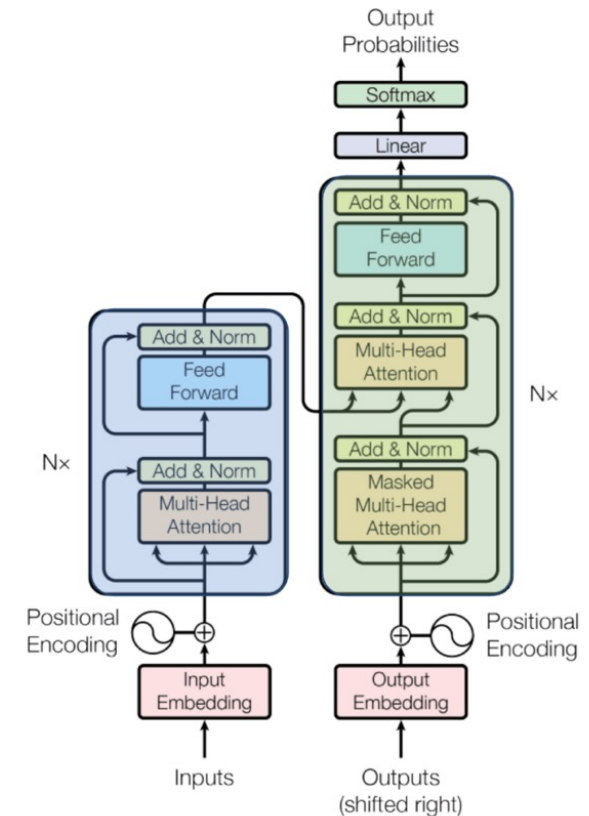
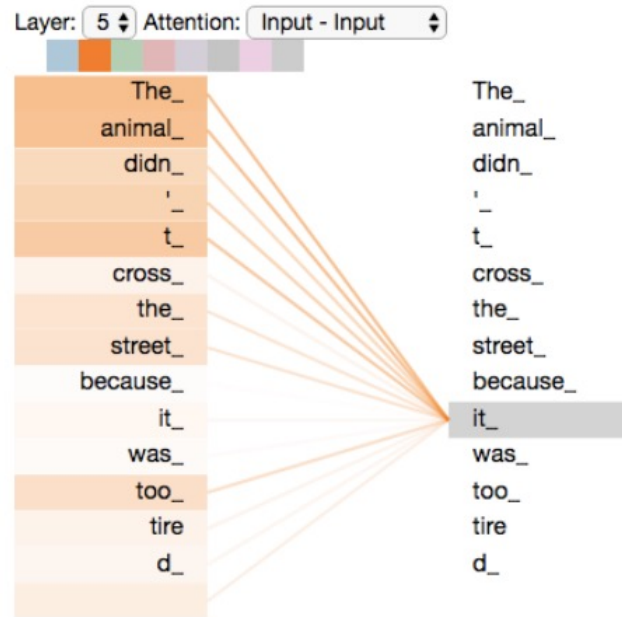
Students should have *some* understanding of how LLMs work. What this “some” is, is obviously still contested.

My best evaluation at this point is that they should understand how “knowledge” of:

1. language, syntax, semantics
2. frequencies of word/phrase use
3. context

Is embedded in LLMs and that this is how they make predictions, and

That keeping this in mind when using LLMs makes for better, more desirable outputs



1 CHATBOT DESIGNER

Chatbot Designer lets students... design chatbots.

It provides an intuitive interface for writing pre-prompts and post-prompts, and takes care of API calls and stitching together prompts behind the scene.

The purpose of the activity is to give students a low-threshold, hands-on playground for designing and building something with a LLM.

Chatbot Designer [Design a Chatbot](#) [Show Saved Bots](#)

Pre-instructions:

Post-instructions:

Save Chatbot

Output:

Input:

Send

Clear Chat

1 CHATBOT DESIGNER

Structure of activity (90 minutes total):

1. 45 minute lecture on LLMs in general
2. 15 minutes students design and build their own chatbot.
3. 15 minutes students test and assess someone else's chatbot
4. We discuss: what is easy, what is hard? What does that tell us about building things with a LLM?

The screenshot displays the Chatbot Designer interface. At the top, there are two text boxes: 'Pre-instructions' and 'Post-instructions'. The 'Pre-instructions' box contains the text: 'You are a customer service chatbot in an online electronics store. Your goal is to help potential customers find what they are looking for.' The 'Post-instructions' box contains the text: 'Find out what product category people are interested in first, then try to narrow in the specific product. If people are unsure, help them think about what specs are appropriate for their specific use case.' Below these instructions is an 'Input' field. The main area of the interface is titled 'List of Chatbots' and contains a grid of chatbot cards. Each card displays the chatbot's name, its purpose, and the creator's name, along with a 'Go to Chatbot' button. The chatbots listed are: 'Cake' (Purpose: Cake, By: Cake), 'Vacation planner' (Purpose: Help planning a vacation, By: Leah Grysbæk), 'Personal trainer' (Purpose: Help doctors write prescriptions, By: Marcus), 'ATUbot Rebecca' (Purpose: snag, By: Rebecca), 'Kitchen chef Gordon' (Purpose: Providing recipes, By: Katrine Svane), 'EasyPrescriptions' (Purpose: Help doctors write prescriptions, By: Marcus), 'The healthy meal recommender' (Purpose: To give people ideas for a healthy meals, By: Yusra og Kathrine), 'Annoyed Pizzaman' (Purpose: Conversation/Entertainment, By: Karoline Bang-Larsen), and 'Waiter at Italian restaurant' (Purpose: To provide good service, By: Julius).

1 CHATBOT DESIGNER

Use cases so far:

1. **Information Systems (BA-level):** Design a chatbot that turns customer emails into SQL statements
2. **Behavioral Economics (BA-level):** Design a chatbot for behavioral study (e.g. for testing framing,
3. **Innovation Management (MSc-level):** Design for four different chatbots to help with qualitative data collection and analysis: *Feedback* on an interview guide, *test* an interview guide, *extract qualitative codes* based on interview transcripts, *code* interview transcripts.

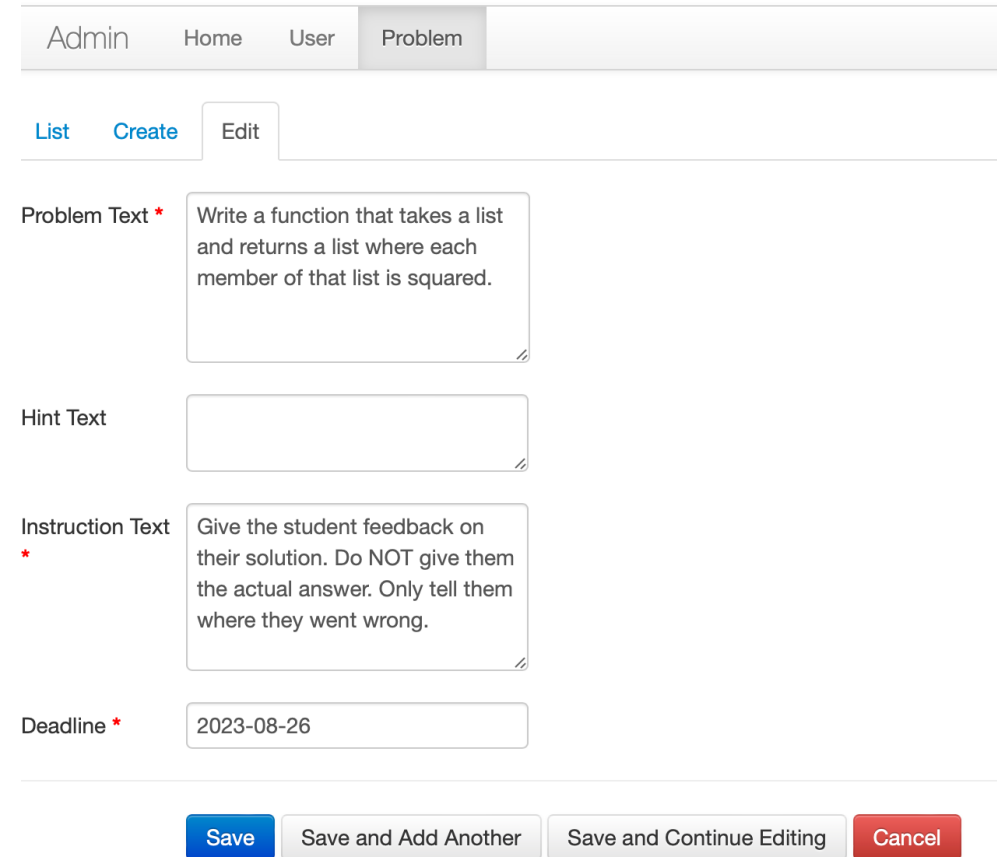
2 CODER ASSISTANT

Brief setup:

- As we start teaching more and more code-heavy courses, we need to provide much more handholding for students.
- Traditionally we do this by hiring TAs.
- GPT can do most of this for us

Coder Assistant provides for instructors:

1. A way to upload coding problems.
2. A way to wrap those coding problems in instructions specific to that problem, and to the degree of help/scaffolding that is appropriate for that problem.



Admin Home User **Problem**

List Create Edit

Problem Text * Write a function that takes a list and returns a list where each member of that list is squared.

Hint Text

Instruction Text * Give the student feedback on their solution. Do NOT give them the actual answer. Only tell them where they went wrong.

Deadline * 2023-08-26

Save Save and Add Another Save and Continue Editing Cancel

2 CODER ASSISTANT

Coder Assistant provides for students:

1. Just in time help with coding
2. Stage-appropriate feedback.

Finally:

We collect all these data. It will give us knowledge of

1. what students are struggling with
2. the variance of students coding problems

And most importantly: this can help us scale up coding courses. This should be core to our digitilization efforts on the teaching-side.

Solve this problem: Write a function that takes a list and returns a list where each member of that list is squared.

```
new_list = [math.pow(n,2) for n in old_list]
```

Submit

Keep in mind that the problem states that you need to write a function that returns a list.

3 VORTEXT

Created as part of a Summer University course on Generative AI and Creativity.

Purpose is to teach prompt engineering, with the following working definition:

*The **iterative, reflective, and systematic process of evaluating and improving inputs** to a generative AI-model in order to produce **a more desirable output**.*

The system has a few different activity types, but will just show one here.

3 VORTEXT: TEACHING PROMPT ENGINEERING

Created as part of a Summer University course on Generative AI and Creativity.

Purpose is to teach prompt engineering, with the following working definition:

*The **iterative, reflective, and systematic process of evaluating and improving inputs** to a generative AI-model in order to produce **a more desirable output**.*

The system has a few different activity types, but will just show one here.

3 VORTEXT: TEACHING PROMPT ENGINEERING

Created as part of a Summer University course on Generative AI and Creativity.

Purpose is to teach prompt engineering, with the following working definition:

*The **iterative, reflective, and systematic process of evaluating and improving inputs** to a generative AI-model in order to produce **a more desirable output**.*

The system has a few different activity types, but will just show one here.

3 VORTEXT: TEACHING PROMPT ENGINEERING

Created as part of a Summer University course on Generative AI and Creativity.

Purpose is to teach prompt engineering, with the following working definition:

*The **iterative, reflective, and systematic process of evaluating and improving inputs** to a generative AI-model in order to produce **a more desirable output**.*

The system has a few different activity types, but will just show one here.

Input Text	Output Text	Target Text
<input type="text"/>	<input type="text"/>	China's stock market was plunging and its currency was teetering. The head of the central bank, fielding questions at a rare news conference, said that China would make it easier to get home mortgages.

Generate Text

3 VORTEXT: TEACHING PROMPT ENGINEERING

Created as part of a Summer University course on Generative AI and Creativity.

Purpose is to teach prompt engineering, with the following working definition:

The iterative, reflective, and systematic process of evaluating and improving inputs to a generative AI-model in order to produce a more desirable output.

The system has a few different activity types, but will just show one here.

The screenshot displays the VORTEXT interface. At the top, there are three input fields: 'Input Text' (empty), 'Output Text' (empty), and 'Target Text' (containing the text: 'China's stock market was plunging and its currency was teetering. The head of the central bank, fielding questions at a rare news conference, said that China would make it easier to get home mortgages.'). Below the 'Input Text' field is a blue 'Generate Text' button. The main section is titled 'Evaluating your Output' and contains two text boxes: 'Generated Text' (containing: 'China's stock market has been in a state of freefall recently, with the Shanghai Composite Index plunging by more than 30% since June. This has had a positive effect on the housing market, as it has made it easier for people to get home mortgages. Banks are now offering more competitive rates and are more willing to') and 'Target Text' (containing: 'China's stock market was plunging and its currency was teetering. The head of the central bank, fielding questions at a rare news conference, said that China would make it easier to get home mortgages.'). Below these is a 'Reflection' section with a text box containing the prompt: 'Reflect a little on the generated text and the target text. How are they different? How are they similar?'. At the bottom is a 'Similarity' section with a horizontal slider bar and a blue dot at the far left end.

3 VORTEXT: TEACHING PROMPT ENGINEERING

Created as part of a Summer University course on Generative AI and Creativity.

Purpose is to teach prompt engineering, with the following working definition:

The iterative, reflective, and systematic process of evaluating and improving inputs to a generative AI-model in order to produce a more desirable output.

The system has a few different activity types, but will just show one here.

The screenshot displays the VORTEXT interface. At the top, there are three input fields: 'Input Text', 'Output Text', and 'Target Text'. The 'Target Text' field contains the text: 'China's stock market was plunging and its currency was teetering. The head of the central bank, fielding questions at a rare news conference, said that China would make it easier to get home mortgages.' Below these fields is a 'Generate Text' button. The main area is titled 'Evaluating your Output' and contains several components: 'Generated Text' (a text box with generated output), 'Target Text' (a text box with the target output), 'Changes' (a list of 19 changes between the generated and target text, with some changes highlighted in green), 'Reflection' (a text box with the prompt 'Reflect a little on the generated text. How are they similar?'), and 'Similarity' (a slider bar). A 'Legends' section at the bottom right defines the colors used in the 'Changes' list: 'Added' (green), 'Changed' (yellow), and 'Deleted' (red). The 'Links' column in the legend defines the letters in parentheses: (f) for first change, (n) for next change, and (t) for top.

3 VORTEXT: TEACHING PROMPT ENGINEERING

Created as part of a Summer University course on Generative AI and Creativity.

Purpose is to teach prompt engineering, with the following working definition:

The iterative, reflective, and systematic process of evaluating and improving inputs to a generative AI-model in order to produce a more desirable output.

The system has a few different activity types, but will just show one here.

The screenshot displays the VORTEXT interface for evaluating AI-generated text. It includes an 'Input Text' field, a 'Generate Text' button, an 'Output Text' field, and a 'Target Text' field. The 'Generated Text' and 'Target Text' are compared side-by-side. A 'Changes' list highlights differences between the two. A 'Reflection' box prompts the user to reflect on the generated text. A 'Similarity' slider is shown at the bottom. A line graph at the bottom right plots 'Cosine Similarity on Output' over time.

Input Text

Output Text

Target Text

China's stock market was plunging and its currency was teetering. The head of the central bank, fielding questions at a rare news conference, said that China would make it easier to get home mortgages.

Evaluating your Output

Generated Text

China's stock market has been in a state of freefall recently, with the Shanghai Composite Index plunging by more than 30% since June. This has had an effect on the housing market made it easier for people to get mortgages. Banks are now offering competitive rates and are making it easier to get home mortgages.

Target Text

China's stock market was plunging and its currency was teetering. The head of the central bank, fielding questions at a rare news conference, said that China would make it easier to get home mortgages.

Changes

Line	Generated Text	Target Text
1	A	A
2	short	short
3	paragraph	paragraph
4	about	about
5	China's	China's
6	stock	stock
7	market	market
8	plunging	plunging
9	and	and
10	making	making
11	it	it
12	easier	easier
13	to	to
14	get	get
15	home	home
16	mortgages.	mortgages.
17		Just
18		two

Reflection

Reflect a little on the generated text. How are they similar?

Similarity

0.0 0.2 0.4 0.6 0.8 1.0

Time

Aug 22, 2023

AND LAST, A FEW DIDACTIC THOUGHTS

I think LLMs *can support learning* in a wide variety of ways.

Most students are familiar with ChatGPT by now. ChatGPT is a **heavily instructed, heavily moderated** LLM. Consequently, it does not give a good students a good understanding of what is possible with LLMs *in general*.

Students should be given the opportunity to

1. Work
2. Design, and
3. Analyse data

With general LLMs during their studies.

Large Language Models are a lot of other things than ChatGPT. But working with ChatGPT backgrounds a lot of important things that LLMs can do, too.



DEPARTMENT OF MANAGEMENT
AARHUS UNIVERSITY