Large Language Models, CIS 1100, & You

Attribution: Today's slides adapted (& sometimes borrowed directly) from CCB's CIS 4210 slides.

Language Models

At its core, a **language model** is a program that estimates the probability of a certain word appearing within a sequence of other words (called the **context**).

LLMS

Language Models: Fill in the Blank

"CIS 1100 is a really _____ class."

word	estimated probability
hard	0.2
fun	0.7
easy	0.1

Can use the probabilities to generate text—fill in the blank with the most likely word.

Language Models: Open-Ended Text Generation

If we take the idea of "fill in the blank" and apply it to the case where the blank comes at the end of a sentence, we can generate text.

- Generate the most likely word given the previous context

- Then, add the word to the context and repeat

Language Models: Open-Ended Text Generation

"How are you?"

___

Language Models: Open-Ended Text Generation

"How are you?"

"I ____"

Language Models: Open-Ended Text Generation

"How are you?"

"I am ___ "

Language Models: Open-Ended Text Generation

"How are you?"

"I am well____"

Language Models: Open-Ended Text Generation

"How are you?"

"I am well. <STOP>"

9

Neural Language Models & Large Language Models

Neural Language Models are those that are built using neural networks.

Large Language Models are those that are built on huge training sets and contain many *parameters*, which are the values that the neural network learns.

Examples:

- GPT-4, GPT-o (from OpenAI)

- Claude 3.5 (from Anthropic)

- LLAMA (from Meta)

- BERT (from Google)

How to Build a Large Language Model

- A little bit of math

- A lot of data

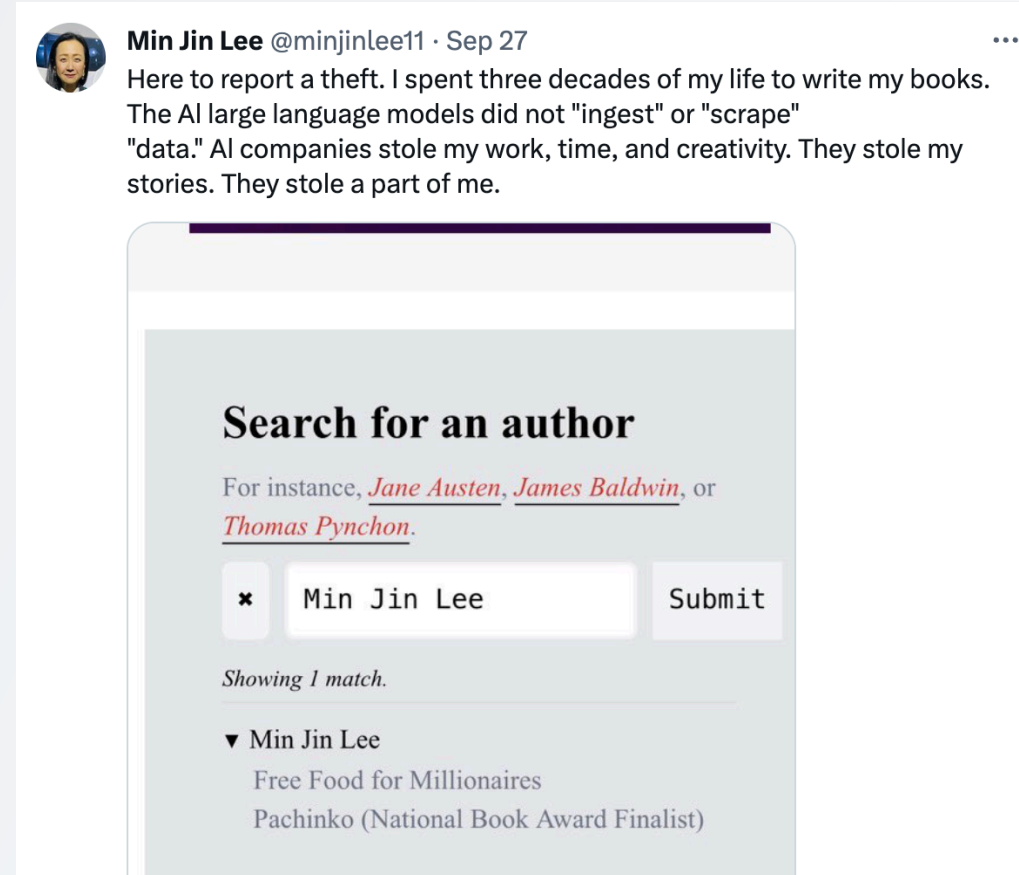- Unbelievable amounts of money and energy—including nuclear energy!

How to Build a Large Language Model

Inputs:

- Webpages & posts on the internet

- Books (including copyrighted ones)

- Synthetic data

Outputs:

- A very complicated mapping of words to *vectors*

- A very complicated mapping from a context to a probability distribution of words that fit

Min Jin Lee @minjinlee11 · Sep 27

Here to report a theft. I spent three decades of my life to write my books. The AI large language models did not "ingest" or "scrape" "data." AI companies stole my work, time, and creativity. They stole my stories. They stole a part of me.

Search for an author

For instance, *Jane Austen*, *James Baldwin*, or *Thomas Pynchon*.

✕ Min Jin Lee   Submit

Showing 1 match.

▼ Min Jin Lee
Free Food for Millionaires
Pachinko (National Book Award Finalist)

Words & Vectors

- Need to do math to do estimation of probabilities

- Can't do math with words!
    - `0.4 * "hard" = ???`

- Instead, we learn multidimensional representations of words that are called *vectors*

- Foundational idea of computational linguistics: similar words are used in similar contexts

- Desirable properties of learned vectors: similar words have similar vectors
    - `[0.1, 0.3]` and `[0.11, 0.28]` are "similar" vectors, or at least more similar than compared to `[40, 50]`

The "Byproduct": Word Embeddings

An essential part of a neural language model is its learned mapping of words to vectors called a **word embedding**.

- This embedding is a way of representing text as numbers

- Good embeddings encode semantic meaning of text by grouping the vectors of similar words together

- Can visualize this a bit here.

Word Embeddings & Vectors

Since words get mapped to vectors, we can do math with them!

- Subtracting vectors can give us a sense of the "difference" between two words, which can be helpful for doing *analogy problems*

"King" is to "man" as "Queen" is to ...

Word Embeddings (Toy Example)

If we generate a 3D word embeddding, we would have 3 element vectors to represent each word.

Word One	Embedding	Word Two	Embedding
Sleep	(0.286, 0.331, 0.087)	Sleeping	(0.306, 0.039, 0.605)
Run	(0.385, 0.155, 0.192)	Running	(0.404, -0.129, 0.711)
Try	(0.779, 0.994, 0.167)	Trying	(0.78, 0.707, 0.667)

Going from the present form of the verb to the participle is typically an offset of about (0, -0.3, 0.5)—this is the *participlifying* vector!

16

Word Embeddings & Bias

Implicit biases are unconscious associations between certain concepts and attributes.

- e.g. if the concept is "flowers", most people automatically associate it with "pleasant" or "beautiful"

- We have ways of studying and measuring implicit bias, going back to landmark sociological studies in the '90s.

- You can take the Implicit Association Test for yourself, too.

Word Embeddings & Bias

Some findings on US participants' implicit biases:

- "European-American names" associated readily with pleasant words, "African-American names" associated with unpleasant words

- "Male names" & math and science go together, "Female names" & the arts go together.

Unfortunately, the LLMs learn from data with all the same biases from human-generated text. Fortunately, de-biasing techniques do exist.

LLMs & Coding

These things are **really** good at reading and writing code. Not perfect, but really good.

It's absolutely true that...

- programmers use LLMs (ChatGPT, GitHub Copilot) to help them read & write code
- it's important for students to learn how to use these tools

A caveat: it's very important that you understand the limitations of the tools in addition to their capabilities.

LLMs & Coding & Penn

These lectures are absolutely **not** an endorsement of you using ChatGPT or Copilot on your assignments in this class, or CIS 1200, or CIS 2400, or...

- You always need to follow the course's policies on use of LLMs.

- It's absolutely vital that you understand how to write code for yourself so that you don't rely on an outside source to be fluent

- Somnetimes these tools are **just totally wrong** and you need to be able to verify when they're right vs. almost right.

LLMs are Confidently Wrong

```
## How to Build a Large Language Model
Inputs:
- Webpages & posts on the internet
- Books ([including copyrighted ones](https://www.
theatlantic.com/technology/archive/2023/09/
books3-database-generative-ai-training-copyright-infrin
gement/675363/))
- Synthetic data? (e.g. [ChatGPT](https://arxiv.org/
abs/2110.09725))
```

This last line was completed for me using GitHub Copilot. Totally unprompted and without caveat.

LLMs are Confidently Wrong

https://arxiv.org/abs/2110.09725

arXiv > gr-qc > arXiv:2110.09725

General Relativity and Quantum Cosmology

[Submitted on 19 Oct 2021]

Multipole moment and singular source in Newtonian gravity and in Einstein gravity

Yu-Zhu Chen, Yu-Jie Chen, Shi-Lin Li, Wu-Sheng Dai

The multipole moments are defined as the multipole expansion coefficients of the gravitational field at infinity. In Newtonian gravity, the multipole moments are determined by the source distribution -- the multipole integrals of the source. In this paper, we show that the multipole moments in general relativity cannot be determined by the multipole integrals of the source. We provide the multipole integrals in static axial spacetimes, such as, the Curzon spacetime. The Curzon spacetime possesses the same multipole integrals of the source with the Schwarzschild spacetime, while they possess different multipole moments.

????????

LLMs are Confidently Wrong

I use LLMs all the time now to help me develop course materials. They're great for generating examples, but I always have to check them.

> **H** Give me a list of 10 fruits starting with a, b, c, d, and so on
>
> Sure, here you go:
>
> A - Apple
> B - Banana
> C - Cherry
> D - Date
> E - Elderberry
> F - Fig
> G - Grape
> H - Honeydew Melon
> I - Kiwi
> J - Jackfruit

24

LLMs are Confidently Wrong

> **H** Is there no fruit that starts with I?

> **⊕** I apologize for the oversight. Let's replace "Elderberry" with "Indian Fig" to include a fruit that starts with I. Here's the corrected Java String[] literal:
>
> ```java
> java                                                    Copy code
>
> String[] fruits = {"Apple", "Banana", "Cherry", "Date", "Indian Fig", "Fig",
> ```

???????????

25

ChatGPT & You

LLMs for Understanding

Accepting the caveat that they can tell you something **wrong**, it could be useful to try to seek alternative explanations for class concepts.

Ask:

- Does this answer sound familiar to something I've heard already?

- Can I use this answer to explain an example I've seen before?

You
I don't understand the point of interfaces in Java.

ChatGPT
In Java, an interface is a way to achieve abstraction, which is a fundamental concept in object-oriented programming. Abstraction allows you to define the structure of a class without providing the implementation details. Interfaces provide a mechanism for defining a contract that a class must adhere to, without specifying how the class should implement that contract.

Here are some key points about interfaces in Java:

LLMs for Understanding

Although CIS 1100 course materials were used to train GPT-3.5, it's not a TA for the course. It won't always understand why you're asking certain parts of a question.

H **You**
OK, that makes sense. Now, I'm having trouble understanding Nodes in Java. Why does setting head = new Node() give me a Tour that has only one node?

ChatGPT
It seems like there might be a bit of confusion here. The code `head = new Node();` doesn't necessarily create a "Tour" with only one node; rather, it creates a new node and assigns it to the `head` reference. The concept of a "Tour" is not explicitly mentioned in your code snippet, so I'll provide a general explanation regarding nodes in the context of linked lists, which is a common use case for nodes in Java.

LLMs for Understanding

Nevertheless, the answer is pretty helpful...

If you want to create a linked list with multiple nodes, you would need to create additional nodes and link them together. For example:

```java
// Creating nodes
Node head = new Node(1);
Node second = new Node(2);
Node third = new Node(3);

// Linking nodes
head.next = second;
second.next = third;
```

In this example, you have a linked list with three nodes, and the `head` reference points to the first node. Each node has a `next` reference pointing to the next node in the sequence.

LLMs for Understanding

LLMs are often good at reading your code and even explaining why it's not doing what you think it should be doing.

You
OK, so can you help me see why my "insertAfter" function is not working correctly?

```
public static void insertAfter(Node n, int data) {
    n.next = new Node(data);
    toInsert.next = n.next;
}
```

ChatGPT
It looks like there might be a bit of confusion in your `insertAfter` function. If you want to insert a new node after a given node (`n`), you need to perform the following steps:

1. Create a new node with the given data.
2. Set the `next` reference of the new node to the `next` reference of the given node.
3. Update the `next` reference of the given node to point to the new node.

Too Much Information

But, unfortunately, they're designed to be as helpful as possible, so it's hard to get them to stop before giving you the actual solution...

Here's a modified version of your `insertAfter` function:

```java
public static void insertAfter(Node n, int data) {
    if (n != null) {
        Node toInsert = new Node(data);  // Step 1:
        toInsert.next = n.next;          // Step 2:
        n.next = toInsert;               // Step 3:
    }
}
```

Too Much Information

The solution on the previous slide is correct; unfortunately, it's also just the solution to the problem itself. This raises **two** significant and incontrovertible issues.

1. You're learning so much less as a beginner programmer if the solution is revealed to you.

2. Submitting work that you did not generate as though it were your own work is a violation of academic integrity & plagiarism policies for both this course and the university overall.

Some Teaching Philosophy for You

Points 1 & 2 on the previous slide are intimately related

- Assignments are designed so that you both develop & demonstrate understanding while doing them.

- The hope is that you are motivated (& have the time and resources) to tackle these assignments on your own, then.

- The policies & codes of integrity are "cudgels" that attempt to insist this on you.

Putting these together, I am trying to provide some **extrinsic** motivation for you to do the work on your own while trying to help foster your own **instrinsic** motivation for the same end.

Some Teaching Philosophy for You

- My goal is to help you learn as much as possible about programming and help you feel confident in your abilities.

- Your goals are probably multifaceted: you want to learn, you want to fulfill a requirement, you want to get a good grade, you want to get a good job, you want to have fun & personally enriching time in undergrad, etc. etc. etc.

Working Through the Design Process

1. Understand the Problem

2. Formalize the Interface

     i. can be literal interfaces, or just some way of specifying how the program will

       be interacted with

3. Write Test Cases

4. Implement the Required Behavior

Pair Programming

A common software development technique involving two participants.

- One person types the code into the computer

- The other reviews each line of code as it's generated

Helpful for spotting bugs, mutually verifying understanding, etc.

ChatGPT as a Pair Programmer

The "Tier List" Problem (sp22)

Help CIS 1100 design some objects for representing tier lists. We'll use two classes to do this: Tier.java and TierList.java. A Tier is like a node: it stores some data (a String[] representing the entries at this Tier) and a pointer to the next Tier. A TierList is a linked sequence of Tier objects.

Way Too Eager

What happens if you try something like this?

> **H** **You**
> I want to implement a Tier List class in Java. Can we work together step by step to design this and program it?

Way Too Eager

39

In this step, we've created a simple `Tier` class with a `name` attribute. The `Tier` class represents a tier in your Tier List.

Step 2: Define the Item class

```java
public class Item {
    private String name;

    public Item(String name) {
        this.name = name;
    }

    public String getName() {
        return name;
    }
}
```

Here, we've created an `Item` class. An item is what you will be ranking in your Tier List.

Step 3: Define the TierList class

```java
import java.util.HashMap;
import java.util.Map;

public class TierList {
    private Map<Tier, String> tiers;

    public TierList() {
        tiers = new HashMap<>();
    }

    public void addTier(Tier tier, String description) {
        tiers.put(tier, description);
```

Rein it in...

> **H** **You**
>
> We're going to work together on a pair programming exercise. I need you to assume the role of a programmer on a team. It's important that we work through things incrementally, and I need you to refrain from answering more than I ask at any point.

Some Things to Keep in Mind

1. Stick to the script!

    i. Make sure to follow the design process! It can feel tedious when you are writing code on your own, but it's vital if you're communicating with another person or agent

2. Go slowly and ask very specific questions/provide very specific directives

3. You can always ask for clarification, more information, or a different approach

4. Make sure to check the output of the program

    i. Read it to make sure you understand what it's suggesting

    ii. COMPILE!!!