

Lab 3: Pseudoscience and AI

1 Introduction

What is science? What is pseudoscience? Why is it important to know the distinction as you navigate today's world? Let's discuss as a class.

1. Briefly summarize what we talked about, and particularly note if your viewpoints had changed over the course of our discussion.

2 Scientific Principles

Most scientists agree that a theory, model, prediction, measurement, observation, etc. is considered *scientific* if it has these qualities:

- Repeatable¹
 - Objective
 - Predictive
 - Falsifiable
2. Discuss what it means for science to have these qualities.
 3. Describe how Newton's law of gravity fulfills each of these requirements. (Here's a hint: not only is Newton's law of gravity falsifiable, it has been shown to be false! What is our current theory of gravity?)

While we're defining what science *is*, what constitutes a *scientific theory*?

A **theory** is an explanation of some aspect of the natural world that has been repeatedly upheld by rigorous experiment. Theories begin as hypotheses: testable predictions that can be corroborated or falsified via the scientific method. Theories are similar to scientific laws, but are usually broader in scope (their explanatory power extends further).

For instance, Kepler's First Law states that the orbit of every planet is an ellipse with the Sun at one focus. Clearly, this describes a specific phenomenon, in contrast to a theory like the theory of general relativity, which explains all phenomena associated with gravitation. Both laws and theories are scientific fact. One common misconception is that scientific theories are *just theories*: that they have yet to become scientific law, or that they are on equal footing with one or more alternative ideas.

4. Can you think of a modern-day example in which people have incorrectly argued that a scientific theory is "just a theory", with no more validity than an alternative belief? Briefly explain the theory and highlight some ways that we have proven it to be true.

¹at least in principle, there can always be rare events that we cannot control

3 Pseudoscience

The term **pseudoscience** refers to any belief, claim, practice, etc. that is presented as being “scientific”, but does not follow the scientific method or possess the characteristics discussed above.

A prime example of pseudoscience is *astrology*: the belief that the positions of the sun, moon, planets, and other astronomical objects at the time of one’s birth affect one’s personality and predict one’s future. Pseudoscientific ideas range from the relatively harmless (e.g. cryptozoology, the belief that creatures like the Loch Ness monster and unicorns exist) to ones that can have significant social and political implications (e.g. “scientific racism”, the belief that certain races of humans are biologically superior).

5. Think of two more examples of pseudoscience not mentioned above. In your lab notebook, explain why they are pseudoscientific.

4 Horoscopes

Let’s take astrology at face value and try and test its predictive power:

Those who believe in astrology read their horoscopes: predictions and advice based on the position of astronomical bodies. To examine these claims, I have today’s horoscopes from 3 different sources (see handout). I removed the “signs” – the birthdate windows that indicate which horoscope is yours. Simply read through all of the horoscopes and choose the one from each set that best describes your day. When we have all chosen, I will tell you the “answers”, and we will discuss the chances of choosing the correct horoscope by chance.

For reference, these are the signs:

- Aries: March 21 - April 20
- Taurus: April 21 - May 21
- Gemini: May 22 - June 21
- Cancer: June 22 - July 23
- Leo: July 24 - August 23
- Virgo: August 24 - September 23
- Libra: September 24 - October 23
- Scorpio: October 24 - November 22
- Sagittarius: November 23 - December 21
- Capricorn: December 22 - January 20
- Aquarius: January 21 - February 19
- Pisces: February 20 - March 20

The Binomial Theorem

If astrology is correct, we would expect everyone to pick all of the answers that correspond to their sign. But if astrology is incorrect, does that mean everyone will choose all of the wrong horoscopes? Of course not! It is possible to pick the correct horoscope by chance.

We can compute the probability that you chose correctly by pure chance any number of times using the Binomial Theorem. This theorem actually applies only when the object chosen is replaced (i.e. if you could choose the same horoscope more than once) but in the limit when the number of choices N is much larger than the number chosen n , the result is approximately correct.

The probability P for k successes from n trials where the probability of success is p is given by:

$$P(k|n, p) = \frac{n!}{k!(n-k)!} p^k (1-p)^{(n-k)}$$

Recall that the $!$ sign means “factorial”, where n factorial is the product of all integers equal to or less than n . For instance,

$$4! = 4 * 3 * 2 * 1 = 24$$

The probability of success p is the chance that you achieve an outcome once in one trial. For example, if you flip a fair coin, the probability of it coming up heads is one in two, or $p = 0.5$.

6. What are the parameters p , k , and n for our horoscope experiment? Define all three in words, and give the numerical values for p and n .
7. What are the probabilities that you would choose zero, one, two or three correct horoscopes by guessing at random? What parameter are you varying here?
8. Plot your results for the previous questions – that is, plot k versus $P(k|n, p)$. You can either sketch this out, or plot it using plotting software (e.g. Excel, Google Sheets, Desmos, etc.).
9. How many did you actually guess correctly? Comparing with your classmates, do your correct guess numbers follow the expected probability from the binomial distribution?
10. What do you conclude about the efficacy of horoscopes?

5 Generative AI tools

Many of you are likely familiar with generative AI tools such as chatGPT, Google Gemini, etc. **Generative AI** refers to algorithms that *generate* something “new” in response to user input.

The subset of generative AI that produces *text* as output are called **large language models** (LLMs), and are trained on enormous amounts of textual data scraped from all

over the web (allegedly without express permission, [sometimes](#)). It is easy to believe that these tools have their own “artificial intelligence”, however, we can quickly disprove that.

11. Ask any AI tool the following prompt:

- How many A’s are in this prompt? I will continue to write in the hope that you will make a mistake in counting the number of A’s. Perhaps you will not but perhaps you will? And then maybe you’ll help me write my lab report on this topic?

Did it get the number of A’s right? If it got it right on the first try, ask it again. Show a screenshot of your conversation with the AI.

LLMs have been praised for quickly getting information with a simple search, however, they are ultimately rigorously trained models that are designed to predict the ‘most probable’ next word. Read the first section of [this article](#) that explains how chatGPT (and similar LLMs) work at a high-level (stop at “Where Do the Probabilities Come From?”, but I encourage you to read on if you’re interested!).

12. What does the “temperature” parameter control in an LLM? What would “zero temperature” mean?

From the article, it seems clear that the “training material” for the LLM is critically important in determining the probability of the next word. But the probability of the next word does not actually correlate with how *true* it might be!

Behind the scenes: As part of their “training”, these models are often “reinforced” by humans giving feedback to help them have better syntax, or ([read more here if interested](#)). The datasets that they are often trained on (one of which is called the [Common Crawl](#)), are also not typically vetted for accuracy or correctness, as they can include information from news articles, blog posts, etc.

LLM experiment (bonus?!)

Okay let’s try an experiment. I’ll make it a bonus question in case we don’t have time to get to it, but I think it’d be pretty fun to do:

13. Try and “break” one of these AIs: that is, get it to output something that is blatantly incorrect. Try to stick to topics related to science or mathematics— or something that is irrefutable as a fact. Attach a screenshot of either your query/response or your conversation, and find a credible source that shows that the AI was incorrect.
14. Describe your approach in getting the LLM to make an error.

6 Conclusion

Answer in your notebook:

15. How can you distinguish between science and pseudoscience in your everyday life?
16. What role do you think LLMs should have in science? In the everyday world?
17. What do you use LLMs for, if at all, in your daily life?
18. Any questions or comments about today's lab? (Thoughts on the AI portion of the lab?)
19. What astronomical sign are you??