Rules for Day 2
Group # = Birth day % 10

- Divide in 10 small groups based on the algorithm above (Birth day % 10)
- **Choose one** of these questions depending on your interests
- Discuss for ~ 15 minutes - Take notes (see below) so that we keep a reference
- Choose a spokesperson to report on your conclusions
    - **One idea per group**
    - **One memorable thing about exploring Japan**

Questions

1. Can you see a credible path towards building AI systems with a deep understanding of data? What would you be most excited to use such a system for?

2. If we assume that "inference works" and all that we need to worry about is model misspecification, what strategies could we imagine to be able to derive robust inference ? What is the most exciting outcome of that?

3. What other major challenge do you see moving forward in using AI for physics? Any idea how to address it? What would be the most exciting outcome if such a challenge could be solved?

Notes:

Group 1 - Notes:
- **Question 2:**
    - Misspecified priors can be updated after the fact.
        - We could build forward models to connect model to data (which we may be able to learn)
        - GitHub Copilot can write better simulation models!
    - Maybe we can use summary statistics that are robust to systematics, and iterate with experts. Maybe better physical priors coming from Foundation Models.
    - Expend the prior when you realize your problem is not constrained enough.
    - Building tests for robustness (similar to scale cuts in traditional methods)
    - Memorable about Japan: So safe, so clean, so happy :-)

Group 2 - Notes:
- **Question 1:**
    - Human experts often work with prior knowledge, they are not purely data-driven. On the other hand, traditional AI methods (excluding foundational models) are

too restricted in terms of how much contextual knowledge they have. (For instance – strong lens vs edge on galaxies)
    - Deep understanding of data for physicists ~= ML model does not violate physics at play in the data it trains on. How to check whether its understanding is deep depends on the task / data, but some degree of out-of-distribution generalization is always a good test.
    - Interpretable latent space with forward model
    - Importance of data topology: local correlations matter

Group 3 - Notes:

Group 4 - Notes:

Group 5 - Notes:

Group 6 - Notes:

- Question 1:
    - AI should be working on raw data
    - Physically motivated models
    - Need models with reason, like the ChatGPT scratchpad
    - Most excited to use such a system for:
        - Grant applications 😛
        - Exploration
        - Anomaly detection? Looking for outliers?
    - Memorable things about Japan
        - Bright lights!
        - Confusing train system
        - Short-ranged automatic doors

Group 7 - Notes:

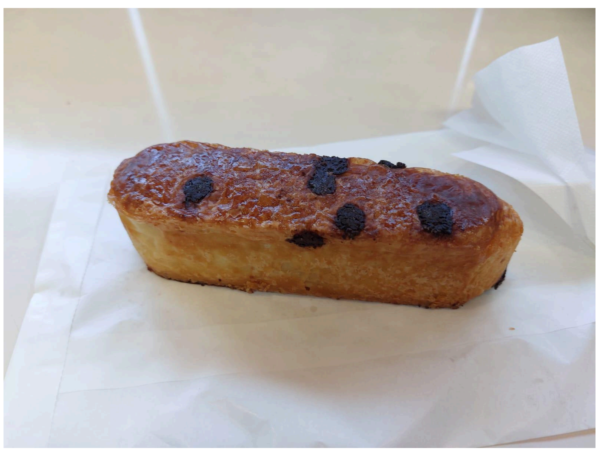
Question 3:

- Major challenge: Convince other people
- Time scale: maybe in years to decades -> but eventually ML should be the norm?
- Moving too fast, hard to trust -> advance in uncertainty quantification
- Some skeptics just don't understand ML, it's not a black box anymore, ML experts should make tutorials to help other people understand
- **Interpretability?**
- **Other challenge: limited computing power -> reproducibility, be open source!**
- **Also beginner-friendliness; even if you're open with your stuff good documentation is necessary and not always there**

Group 8 - Notes:

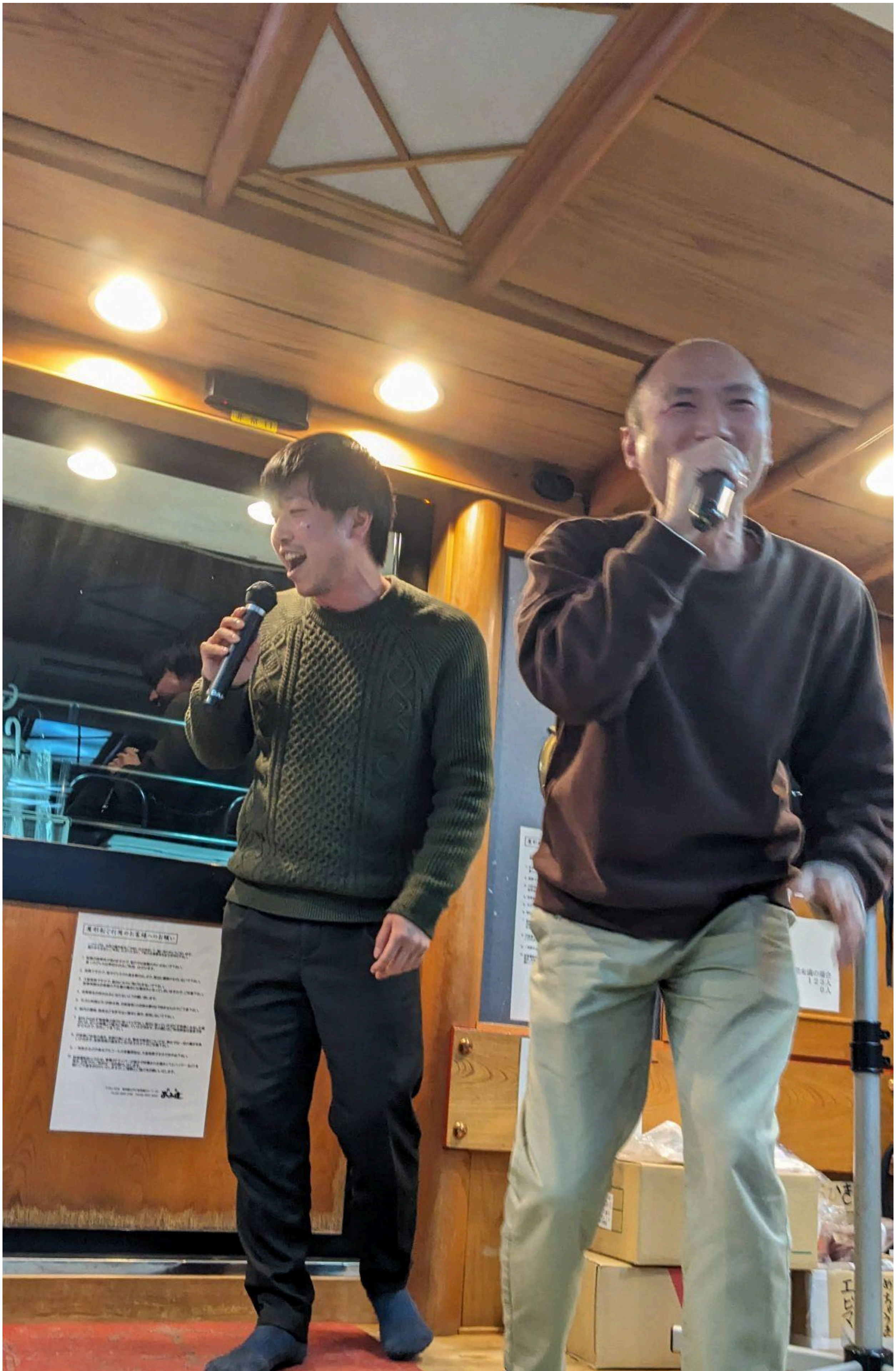
Can you see a credible path towards building AI systems with a deep understanding of data? What would you be most excited to use such a system for?

- What does “understanding data” mean?
    - Need to have confidence about the results
    - Need for interpretability
- No need for human guidance → when to trust it?
-
- Credible path towards building AI: humans should become the caregivers of these machines, providing the requested high-quality data and electricity ($) 🤖
- Excited to use such system: be fully taken care of by AIs (what to eat, where to live, who to marry) and get everything discovered for us (and then take care of the interpretation part when we feel like)



🥐**Japanese Croissant** 🥐

Group 9+10 - Notes:



Day 1

Rules

- Divide in 12 small groups based on your month of birth
- Each group has 1 assigned question (see below) - Each question is discussed by 2 groups
- Discuss for ~ 20 minutes - Take notes (see below) so that we keep a reference
- Choose a spokesperson to report on your conclusions

Questions

1. Are we done training models from scratch in physics? Yes / No / Not yet? How far are we from big foundation models for physics / astro? Main bottlenecks? Multimodal? Data heterogeneity? Compute? Non homogeneous coverage? Will those models change the field of AI4Phys?

2. How do we move from proof of concept to deployment? (Robustness) What are the main key developments needed for robust AI based inference?

3. Will AI be the main tool for discovery with future surveys (Rubin, Euclid etc)? When will this happen and how? Why has it not happened yet with current data?

4. Are we satisfied with current uncertainty quantification methods?

5. Will LLMs change the way scientific analysis is done? Yes / No / No way? How? Replace Coding? Replace Literature search? Propose research directions? Other?

6. What can we (physicists) bring to ML research? Data? Use cases? More? Nothing?

Notes

Group 1 question 1

Are we done training models from scratch in physics? Yes / No / Not yet? How far are we from big foundation models for physics / astro? Main bottlenecks? Multimodal? Data heterogeneity? Compute? Non homogeneous coverage? Will those models change the field of AI4Phys?

Training from scratch
- Easy to do
- Do we have a database that is enough to train a foundation model that generalizes?

Foundation model for science
- The question may be equivalent to asking, “are there broader datasets that are appropriate for our astrophysical tasks of interest, or are there some tasks for which the broader datasets encode inappropriate biases.” In this context the answer seems to be “no.”
- Too soon to tell if they’ll be a sea change for ai4phys

Group 2 - Question 6

7. What can we (physicists) bring to ML research? Data? Use cases? More? Nothing?

- Compared to industry, physicist care a lot more about biases and uncertainty quantification
- Interpretability – ultimately we care about understanding the Universe and nature, and robust interpretations of results are crucial.
- Historically, computer vision/industry algorithms have been adapted to physics, but the trend is slowly changing (with advent of PINNs / customs losses etc)
- Wrto data: Data is NOT cheap (because of experiments/simulations) compared to, say, natural images or web scraped text.
- Benchmarking – industry does extensive benchmarking compared to physicists, and it’s more “organized”. But physicsits need more “expert-level” benchmarking.
- Statistical physics inspired theories for ML .
- Industry is not going to solve problems for us, we have to do this ourselves (data formats, tokenizations)

Group 3 - Question 2

Q:How do we move from proof of concept to deployment? (Robustness) What are the main key developments needed for robust AI based inference?

- Simulations are easy, things work very easily, but when we move to data, we quickly find many effect not in the simulations -> usually doesn't work
    - Eventually, things are solvable, but requires a looooot of work.
    - Expert verification of behavior on real data (a case where we can see the failures)
    -

- Moving from prototype to good code.
    - Large scripts
    - Experiment tracking (weight & biases)
    - Moving from TF1 to TF2 => How to find long term solutions
    - Docker could be an option

- Sometimes model are too simple

- Unexpected "things" in the data

- Difficult to identify uncertainty in the data in SBI when we don't fully understand the signal

Group 4 - Question 5
Will LLMs change the way scientific analysis is done? Yes / No / No way? How? Replace Coding? Replace Literature search? Propose research directions? Other

LLMs have already changed the way of scientific analysis to some extent.
Replace coding for sure, not completely (human supervision is required)
Make Literature search a lot easier (more in the future?)

Scientific research directions: help find interdisciplinary connections or suggest scientific workflows from other domains.

An LLM helped the group pick the speaker.

Poe (gpt-3.5-turbo):
Yes, Language Models like LLMs (Large Language Models) have the potential to significantly impact the way scientific analysis is done. LLMs are capable of understanding and generating human-like text, which opens up new possibilities for scientific research and analysis.
Here are a few ways LLMs can change scientific analysis:

Data analysis: LLMs can assist researchers in analyzing large volumes of data by extracting relevant information, identifying patterns, and generating insights. They can automate data preprocessing tasks, perform statistical analyses, and help in interpreting the results.
Literature review: LLMs can aid scientists in conducting literature reviews more efficiently. They can quickly scan through vast amounts of scientific literature, extract key information, and summarize relevant findings. This can save researchers time and help them stay up-to-date with the latest research.
Hypothesis generation: LLMs can help scientists generate new hypotheses by combining and analyzing existing knowledge from various domains. By processing large amounts of text and identifying connections between different concepts, LLMs can propose novel research directions and hypotheses for further investigation.
Experiment design: LLMs can assist researchers in designing experiments by suggesting relevant variables, controls, and methodologies based on existing scientific knowledge. They can also simulate scenarios and predict possible outcomes, helping scientists optimize their experimental designs.
Collaboration and communication: LLMs can facilitate collaboration among scientists by acting as virtual assistants, providing information, answering questions, and aiding in communication. They can help researchers overcome language barriers and improve knowledge sharing within the scientific community.

However, it's important to note that LLMs are tools that assist scientists and should not replace human expertise and critical thinking. While they can enhance scientific analysis, they are not substitutes for domain knowledge, experimental skills, and the scientific method. Researchers should use LLMs as aids in their work and exercise caution in interpreting their outputs.

Group 5 - Question 3

Will AI be the main tool for discovery with future surveys (Rubin, Euclid etc)? When will this happen and how? Why has it not happened yet with current data?
As tools AI will be very likely used.
When will we believe discoveries? Overlap between Bayesian methods and AI
However, in particle physics AI-based discoveries are very common
What makes these datasets so different?

Group 6 - Question 4

Are we satisfied with current uncertainty quantification methods?

Group 7 - Question 4

Are we satisfied with current uncertainty quantification methods?

Group 8 - Question 3
No. Our current mode of operation is we go for things that we things we can find. Propose that to funding agencies. We are then spending that money to build a big instrument but only using the lowest hanging fruits part of the data. Need to spend more money on theory and methods to advance.

Group 9 - Question 5

“Large language models (LLM) are very large deep learning models that are pre-trained on vast amounts of data. The underlying transformer is a set of neural networks that consist of an encoder and a decoder with self-attention capabilities.”

Will LLMs change the way scientific analysis is done? Yes / No / No way?
Get thing faster, pick command from chat GPT
Time domain astronomy possible to go to parameters other than language

How? Replace Coding? Replace Literature search? Propose research directions? Other?
Replace Literature search? Yes
Propose research directions? If constrained to new paper. Pre-trained suggestion are basically old paper biased.From hack GW paper analysis, clustering
How the LLM know if it is the important problem?
Have fixed method, driven by new data

Group 10 - Question 2

How do we move from proof of concept to deployment? (Robustness) What are the main key developments needed for robust AI based inference?

Proof of concept to deployment:
- Open the data little by little, check along the way that things are making sense (or good enough), using interpretability where possible
- BUT - make sure we use blinding (we don’t want to cheat! And overtrain as scientists)
- Scaling up effectively

Robustness:
- Diversity in forward modeling (e.g. using independent and varied simulation codes) and making sure uncertainties take this into account
- Adding in systematics carefully in our forward modelling
- Taking advantage of your prior knowledge of the data - e.g. symmetries to limit your parameter search space to a better subspace

Group 11 - Question 6

What can we (physicists) bring to ML research? Data? Use cases? More? Nothing?

- Bring new problems
  - Failure cases
  - Lots of data to apply specific statistical models
- Different use cases
  - Uncertainty / more precision
  - More quantitative applications
  - State of the art on a specific problem instead of broad average performance
  - → necessity to develop different methods
- Different problem-solving skills / different perspective approach or mentality
- Interpretability / Neural networks as complex systems to study with physics tools
- More “relaxed” attitude / Quality over quantity / Focus on research results rather than hypes, money, profit,
  - Take more time to do more (careful) research / less product-driven mentality (need to deliver)

Group 12 - Question 1