AI Research Scientist Interview Challenge

Background

An AI research lab is exploring novel approaches to improve the reasoning capabilities and factual accuracy of large language models. Despite recent advances, current LLMs still struggle with complex multi-step reasoning, tend to hallucinate facts, and have difficulty with specialized domain knowledge. The lab is interested in developing innovative techniques to address these limitations.

This challenge simulates a research environment where you'll investigate a specific problem in AI capability and propose novel approaches to advance the state of the art. The focus is on originality of thinking, methodological rigor, and the ability to design well-structured experiments to validate your hypotheses.

Objective

Design and implement a research project that advances the state of the art in one of these areas:

1. Enhancing factual accuracy in language model responses
2. Improving multi-step reasoning capabilities
3. Augmenting domain-specific expertise in a technical field

Your solution should include both theoretical contributions and empirical validation through well-designed experiments. The emphasis is on innovative approaches rather than simply applying existing techniques.

Resources Provided

- **Access to**: Open-source language models (e.g., Llama 2, Mistral)
- **Datasets**: Standard benchmarks for factual knowledge, reasoning, and domain expertise
- **Compute**: Details on available compute resources for experiments
- **Literature**: A curated list of relevant papers in the field
- **Evaluation Framework**: Code for standard evaluation metrics and protocols

Scope of Work

1. Research Proposal Development

- Define a specific research question within the chosen area
- Review relevant literature and identify limitations of current approaches
- Develop hypotheses about potential improvements
- Create a detailed research plan with clear milestones

2. Method Design

- Develop a novel approach to address the identified limitations
- Create formal descriptions of algorithms or techniques
- Consider theoretical foundations and limitations
- Compare with existing methods conceptually

3. Experimental Design

- Design rigorous experiments to test your hypotheses
- Develop or adapt appropriate evaluation metrics
- Create baseline implementations for comparison
- Design ablation studies to isolate the impact of different components

4. Implementation & Evaluation

- Implement your proposed method and baselines
- Conduct comprehensive experiments according to your design
- Perform error analysis to understand failure modes
- Iterate on your approach based on preliminary results

5. Analysis & Future Directions

- Interpret experimental results in the context of your hypotheses
- Analyze limitations and potential improvements
- Propose future research directions
- Discuss potential practical applications

Technical Requirements

- Well-documented research code following scientific computing best practices
- Reproducible experimental setup with clear instructions
- Rigorous statistical analysis of results
- Comprehensive documentation of methodological choices
- Clear presentation of results with appropriate visualizations

Evaluation Criteria

Your research project will be evaluated based on:

- Originality and innovation of the proposed approach
- Methodological rigor and experimental design

- Quality of implementation and analysis
- Clarity of documentation and presentation
- Depth of understanding of the problem space

Discussion Questions

- How does your approach differ from existing methods in the literature?
- What were the most surprising findings from your experiments?
- What are the limitations of your current approach?
- How would you extend this work given additional time and resources?
- What are the broader implications of your research for AI capabilities?

Deliverables

- Complete research paper (6-8 pages in conference format)
- Code repository with implementation and experiments
- Presentation slides summarizing your approach and findings
- Supplementary materials with additional experimental details

Final Notes

This challenge assesses your ability to identify promising research directions, develop innovative approaches, and rigorously evaluate your ideas. We value creativity, scientific rigor, and clear communication of complex ideas. The emphasis is on advancing the field through novel contributions rather than incremental improvements to existing techniques.