Federated Learning for Artificial General Intelligence

MTP Project Proposal

JEYADEV L

M24DE3040

1. Introduction

The pursuit of Artificial General Intelligence (AGI), defined as AI systems capable of matching or surpassing human capabilities across virtually all cognitive tasks [1], represents the pinnacle of AI research. Achieving AGI necessitates overcoming significant hurdles, including the ability to learn from vast and diverse datasets, generalize knowledge across domains, and adapt to novel situations. Simultaneously, the increasing concerns around data privacy, security, and regulatory compliance (e.g., GDPR, CCPA) pose substantial challenges to the traditional centralized data collection and processing paradigms often employed in large-scale AI model training.

Federated Learning (FL) has emerged as a distributed machine learning approach that enables collaborative model training without directly sharing raw data [2]. In FL, local models are trained on decentralized datasets, and only aggregated model updates are shared with a central server. This paradigm offers a compelling solution for privacy-preserving analytics, allowing organizations to leverage collective intelligence while safeguarding sensitive information. The intersection of FL and AGI presents a fascinating research avenue: can FL contribute to the development of AGI by facilitating learning from diverse, distributed, and privacy-sensitive data sources, or can AGI principles enhance FL systems to achieve more generalized and adaptable learning capabilities? This project proposes to explore this intersection, investigating how Federated Learning can contribute to the realization of AGI by addressing data privacy and scalability challenges inherent in training highly capable general AI systems.

2. Problem Statement

The development of Artificial General Intelligence faces several formidable challenges, particularly concerning data management and ethical considerations:

- **Data Scarcity and Diversity for Generalization:** AGI requires learning from an immense diversity of data to achieve broad cognitive capabilities. Centralized collection of such vast and varied datasets is often impractical due to logistical, computational, and legal constraints.

- **Privacy and Security Concerns:** Training AGI models on sensitive real-world data raises significant privacy and security risks. Centralizing such data creates single points of failure and makes compliance with data protection regulations extremely difficult.

- **Scalability of Centralized Training:** The computational resources required to train AGI models on centrally aggregated global datasets are enormous, posing significant scalability challenges.

- **Ethical Implications of Data Usage:** The ethical implications of using vast amounts of personal or proprietary data for AGI development are profound, necessitating privacy-preserving mechanisms.

- **Lack of Generalization in Current AI:** Existing AI models, primarily Artificial Narrow Intelligence (ANI), excel at specific tasks but lack the ability to generalize knowledge or transfer skills across domains without extensive retraining. AGI requires a paradigm shift in how models learn and adapt.

This project aims to address these challenges by investigating how Federated Learning can serve as a foundational component for developing AGI, specifically by enabling privacy-preserving, scalable, and distributed learning from diverse data sources, thereby fostering the generalization capabilities crucial for AGI.

3. Objectives

The main objectives of this project are:

1. To explore the **theoretical foundations** and potential synergies between Federated Learning and Artificial General Intelligence.

2. To design a **conceptual framework** for how Federated Learning can facilitate the development of AGI by enabling **distributed, privacy-preserving**, and continuous learning from diverse data sources.

3. To implement a **prototype Federated Learning system** that demonstrates learning and generalization across multiple simulated or real-world heterogeneous datasets, mimicking aspects of AGI's broad cognitive capabilities.

4. **To evaluate the impact** of Federated Learning on the **generalization ability** and adaptability of AI models, particularly in scenarios where data is distributed and sensitive.

5. To analyze the **challenges and limitations** of applying Federated Learning to AGI development, including scalability, communication efficiency, and convergence in highly heterogeneous environments.

4. Literature Review

Artificial General Intelligence (AGI): AGI represents the goal of AI, aiming to create systems with human-level cognitive abilities across a wide range of tasks [1]. Unlike Artificial Narrow Intelligence (ANI), which excels in specific domains, AGI systems are expected to possess capabilities such as reasoning, planning, learning, natural language understanding, and the ability to integrate these skills to solve novel problems [1]. The development of AGI is a complex endeavour, requiring vast amounts of diverse data and robust learning mechanisms. Ethical considerations, particularly regarding the potential risks and societal impact of AGI, are also a significant area of discussion [1].

Federated Learning (FL): Federated Learning is a distributed machine learning paradigm that enables collaborative model training without centralizing raw data [2]. This approach addresses critical concerns related to data privacy, security, and regulatory compliance. In FL, clients (e.g., individual devices, organizations) train local models on their private datasets, and only model updates (e.g., gradients, weights) are aggregated by a central server to build a global model. This iterative process allows for collective intelligence while keeping sensitive data localized. FL has been successfully applied in various domains, including mobile keyboard prediction, healthcare, and finance, where data privacy is paramount [2].

Intersection of FL and AGI: The concept of Personalized Federated Intelligence (PFI) has recently been introduced as a complement to AGI, focusing on adapting powerful foundational

models (FMs) to meet specific user needs while maintaining privacy and efficiency [3]. PFI integrates the privacy-preserving advantages of FL with the generalization capabilities of FMs, enabling personalized, efficient, and privacy-protective deployment at the edge. This highlights a growing recognition of FL's potential to contribute to more generalized and adaptable AI systems by leveraging distributed data without compromising privacy. However, the direct application of FL to achieve full AGI, which requires a much broader scope of generalization and continuous learning across vastly different domains, remains a nascent research area. This project aims to bridge this gap by exploring how FL can be architected and optimized to foster the generalized learning capabilities essential for AGI, moving beyond personalized applications to truly general intelligence.

5. Methodology

This project will adopt a research-oriented approach, combining theoretical analysis, conceptual framework design, and experimental validation through simulations. The proposed methodology includes the following steps:

1. **Theoretical Analysis and Conceptual Framework Design:**

    o   Conduct an in-depth analysis of the core principles of AGI and the mechanisms of Federated Learning.

    o   Identify key challenges in AGI development that FL can potentially address (e.g., data diversity, privacy, scalability).

    o   Design a conceptual architectural framework illustrating how FL can be integrated into an AGI system to **facilitate distributed, continuous, and privacy-preserving** learning across heterogeneous data sources.

    o   Define the types of data and learning tasks that would be representative of AGI's broad cognitive abilities, including an analysis of the **potential impact of LLM-generated text** on data authenticity and model generalization.

2. **Prototype Federated Learning System Development (Simulation-based):**

    o   Develop a simulation environment that mimics a distributed learning scenario with multiple clients, each holding diverse datasets representing different aspects of general intelligence (e.g., natural language, image recognition, reasoning tasks).

    o   Implement a Federated Learning algorithm (e.g., Federated Averaging) within this simulation environment.

    o   Develop or adapt simple AI models for each client that can learn from their respective local datasets.

3. **Generalization and Adaptability Evaluation:**

    o   Design experiments to assess the generalization capabilities of the FL-trained global model on unseen, diverse tasks that require knowledge transfer across domains.

- o Evaluate the model's adaptability to new, previously unencountered data distributions or tasks within the federated setting.

- o Compare the performance of the FL-trained model against a centralized training approach (if feasible in simulation) and isolated local models.

4. **Scalability and Communication Efficiency Analysis:**

- o Analyze the communication overhead and computational efficiency of the FL framework as the number of clients and data volume increase.

- o Investigate strategies to optimize communication and computation in the context of AGI development.

5. **Privacy Analysis (Conceptual/Qualitative):**

- o Qualitatively assess how the FL framework contributes to data privacy compared to centralized training, considering the nature of shared information (model updates vs. raw data).

- o Discuss potential privacy attacks in FL and conceptual countermeasures relevant to AGI.

6. References

[1] Wikipedia. (n.d.). *Artificial general intelligence*. https://en.wikipedia.org/wiki/Artificial_general_intelligence

[2] IBM Research. (2022, August 24). *What is federated learning?* https://research.ibm.com/blog/what-is-federated-learning

[3] Qiao, Y., Le, H. Q., Raha, A. D., Tran, P.-N., Adhikary, A., Zhang, M., ... & Hong, C. S. (2025). *Towards Artificial General or Personalized Intelligence? A Survey on Foundation Models for Personalized Federated Intelligence*. arXiv preprint arXiv:2505.06907. https://arxiv.org/abs/2505.06907

[4] A Promising Path Towards Auto formalization and General Artificial Intelligence https://doi.org/10.1007/978-3-030-53518-6_1

[5] Federated Learning https://federated.withgoogle.com/#top