Overview

Career Aspirations and Research Vision

As an assistant professor in Mathematics at UAB, I aim to bridge statistical theory with AI through the innovative **Representative Learning** (RL) framework. Developed in response to Federated Learning's (FL) inefficiencies, namely communication overhead, data heterogeneity, and security concerns, RL uses 'representatives' to encapsulate essential data traits from local nodes.

This streamlines central analysis, reduces data movement, and enhances privacy, significantly improving scalability and adaptability. By reducing communication overhead with compact representatives and effectively managing heterogeneous data and tasks, RL ensures privacy, efficiency, and flexibility, setting it apart from traditional FL methods.

My career vision is to refine RL's principles and advocate for its integration into AI and statistics, establishing a new generation of AI systems that are secure, scalable, and grounded in statistical rigor. This effort aligns with my goal to advance AI technology responsibly, enhancing societal benefits.

Educational Philosophy and Objectives

My educational philosophy focuses on crafting a dynamic, research-driven curriculum that blends foundational theories with practical application, equipping students to navigate future technological challenges. This CAREER proposal aims to deeply engage my four PhD students in substantial distributed machine learning projects, enriching their learning and research experiences. I am also dedicated to expanding my mentorship through REU programs and regularly updating course content with the latest research. This strategy aims to develop technically adept and ethically informed professionals poised to lead in AI and statistics. Additionally, the Data Science UAB Summer Camp initiative, bridging the educational gap between high school and university, reinforces our educational objectives by providing essential data science exposure to students and teachers.

Intellectual Merit

My research integrates RL with deep learning, creating scalable, robust frameworks optimized for distributed environments. RL stands out by adeptly handling heterogeneous data and diverse tasks, prioritizing privacy, efficiency, and explainability. Such capabilities make AI systems more understandable and trustworthy, essential for sensitive areas. To actualize this vision, the project sets forth the following primary objectives:

1. **Foundational Expansion of RL:** Adapting RL to accommodate both smooth and non-smooth loss functions, as well as unsupervised learning, broadening its applicability and robustness.
2. **RL in Non-centralized Systems:** Extending RL to decentralized systems, streaming data, and diverse task environments to ensure secure, flexible, and efficient data processing.
3. **Deep Learning Architecture:** The ultimate goal, developing safe and adaptive deep learning frameworks based on RL, leveraging tasks 1, 2 to achieve scalability and robustness in distributed AI.

Achieving these objectives promises to revolutionize AI systems' scalability and reliability, making them more adaptable, secure, and interpretable across diverse and distributed digital ecosystems. Additionally, creating accessible, open-source software in both Python and R will provide the scientific community with advanced tools for distributed learning, fostering greater adoption and innovation in the field.

Broader Impacts

This project amplifies societal benefits by fostering advanced research and educational experiences through high school summer programs and the development of open-source software in both Python and R, enhancing access to state-of-the-art AI tools and encouraging broader participation in STEM. The application of RL in Personalized Medicine and collaborative efforts showcases real-world impacts, advancing public health and promoting scientific collaboration. These initiatives collectively cultivate a diverse, skilled STEM workforce, demonstrating the project's dedication to societal progress.