Research Statement

Pingbang Hu

January 27, 2025

My research focuses on advancing the field of trustworthy machine learning (ML) and data-centric artificial intelligence (AI). The objective of my work is to develop foundational techniques and theories that scale effectively to large-scale AI systems while addressing critical challenges related to transparency, reliability, and accountability. A key focus of my recent research is training data attribution and machine unlearning. These topics are essential for improving our understanding of how data points contribute to model predictions and for enabling systems to comply with privacy demands by provably removing the influence of sensitive data upon request.

In the past year, I have published five preprints, among which three have been accepted at top-tier AI/ML conferences. My publications center around the key problems of training data attribution, machine unlearning, and data curation, highlighting my contributions to advancing the state-of-the-art in trustworthy and scalable data-centric AI research. Among these, two of the notable publications include our open-sourced project on efficient training data attribution library and also a project on interpretable data augmentation from my successful internship at the National Institute of Informatics in Tokyo, Japan. These experiences not only refined my technical expertise but also provided valuable insights into collaborative, cross-cultural research environments.

Besides my research, I am deeply committed to broadening my intellectual horizons. Over the years of my undergraduate and graduate studies, I have engaged with diverse academic disciplines, completing advanced coursework in mathematical statistics, optimization theory, theoretical computer science, and abstract mathematics. This academic journey has equipped me with a solid theoretical foundation while enabling me to address applied challenges with a rigorous approach. These experiences have shaped my holistic view of my research area—as an interdisciplinary field that demands both mathematical rigor and real-world applicability.

My overarching goal is to contribute to a future where AI systems are not only efficient and provable, but also ethical, accountable, and aligned with societal values. To this end, I seek to integrate methodologies from techniques across all related disciplines to create AI systems that are robust, scalable, and trustworthy. My long-term vision is to translate these research advancements into actionable frameworks that influence industry practices and academic understanding alike.

Other Experience and Interests

Before beginning my Ph.D. studies, I explored various research areas, primarily focused on geometry-related theoretical topics. These included graph neural networks, fast graph algorithms (within the context of theoretical computer science), game theory, and deep

learning theory. Overall, I enjoy reasoning in a rigorous, mathematical, and geometric manner, as it often leads to clean, elegant, and intuitive solutions while maintaining the celebrated rigor of mathematics.