DecentraTrain: A Blockchain based Decentralized AI Model Training with Hybrid Cluster Computing

Abstract:

The burgeoning demand for artificial intelligence (AI) and machine learning applications has intensified the need for efficient, scalable, secure, and cost-effective AI model training solutions. Traditional centralized approaches to AI model training face challenges related to transparency, privacy, accountability, and cost. To address these issues, we present DecentraTrain, a novel AI model training platform that combines a hybrid cluster computing model with blockchain technology, staking mechanisms, and enhanced privacy using zero-knowledge proofs (ZKP) to offer a decentralized, incentivized, privacy-preserving, and cost-effective solution for AI model training.

DecentraTrain leverages data and model parallelism within a cluster computing framework to efficiently distribute the computational workload of large-scale AI model training across multiple nodes. This approach reduces the cost of training by utilizing the spare computing capacity of participating users, who can earn rewards for their contributions. The platform employs blockchain technology to create a secure, tamper-proof, and transparent environment for model training. Nodes participating in the training process stake cryptocurrency as collateral, which incentivizes accurate results and deters malicious behavior.

To enhance privacy, DecentraTrain incorporates zero-knowledge proofs in the verification process. ZKPs allow nodes to prove the correctness of their training results without revealing the underlying data or model parameters, protecting sensitive information and ensuring privacy for all participants. This not only maintains the confidentiality of the training data but also prevents potential leakage of proprietary model information.

This paper presents the design, implementation, and evaluation of DecentraTrain, a cost-effective blockchain-based AI model training platform that incorporates a hybrid cluster computing model, staking mechanisms, and enhanced privacy using zero-knowledge proofs. Our results demonstrate the potential of DecentraTrain in terms of scalability, security, privacy, and cost reduction, paving the way for a more democratic, collaborative, and accountable approach to AI model training with enhanced privacy, transparency, and financial incentives for users with spare computing capacity.