Modern humans are currently surrounded by a variety of classical machine learning models offered to them as a service, i.e. not running locally on the user’s device. These models require a lot of computational power to provide people with answers to their questions, e.g. weather forecast or music recommendations. However, even more resources are needed to train such a model, which makes it impossible for the end user to ensure the validity of the training process.

More specifically, how can a classical model owner prove that the model was trained correctly, i.e. according to a public specification and using a specific dataset? In response to this issue a variety of cryptological solutions has emerged recently. Among the first authors to touch upon the topic are Zhao et al. (2021), where they introduce a new general-purpose proof-of-training framework which did not prove itself suitable for use in real-time scenarios. However, in the recent paper Garg et al. (2023) present a novel efficient approach for constructing proofs of training for logistic regression, achieving promising benchmark results on big datasets. Combined with recent advancements in proof-of-inference area for the decision tree model by Zhang et al. (2020), these works have shown that it is possible to solve verifiable machine learning tasks for classical models by making the proving algorithms more specialized on the problem at hand.

At the same time little research has been conducted on the prospects of using metamodels, such as gradient boosting, which proved to be of great interest in production, with big companies such as Yandex and Yahoo! combining classical models with deep neural networks by means of gradient boosting. Such metamodels also typically work faster, and find their usage in scenarios where response time is crucial with the model being constantly partially re-trained.

Research is needed to construct specialized proof-of-training protocols for the decision tree machine learning model. Specifically, shallow decision trees are of great interest, as they are typically used in applications. These protocols in turn should become the base for providing efficient zero-knowledge proof-of-training protocols for gradient boosting and supplementary classical metamodels, such as random forest, which are widely deployed.

We present a novel optimized zero-knowledge framework for verifying gradient boosting training process, based on the proposed shallow decision tree proof-of-training protocol. Benchmarks were conducted which showed applicability in both classical proof-of-training scenarios as well as constant model fine-tuning scenario. Implementation source code is open and available for usage and modification.

The paper is organized as follows. In Section 2 we introduce notation and include background information on zero-knowledge proofs, decision trees and gradient boosting. In Section 3 we present an overview of our techniques. Section 4 describes our proof-of-training protocol for decision trees, which we use in Section 5 to construct proof-of-training protocol for gradient boosting. Implementation and evaluation are presented in Section 6.

REFERENCES:

1. Lingchen Zhao et al. (2021). VeriML: Enabling Integrity Assurances and Fair Payments for Machine Learning as a Service. *IEEE Transactions on Parallel and Distributed Systems, Vol. 32, No. 10, 2524–2540*.
2. Sanjam Garg et al. (2023). Experimenting with Zero-Knowledge Proofs of Training. *In Proceedings of the CCS '23: ACM SIGSAC Conference on Computer and Communications Security, 1880–1894*.
3. Jiaheng Zhang et al. (2020). Zero Knowledge Proofs for Decision Tree Predictions and Accuracy. *In Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS ’20). 2039–2053*.