Auditable Execution of Machine Learning using Blockchain and Trusted Execution Environments

A conceptual framework reconciling *confidentiality* with *transparency*
in data-driven *decision-making processes*
based on *ML outputs.*

N. Romani, V. Gatteschi, R. De Caria, L. Zunino, V. Villa, M. Domaneschi, G. Marengo,
Politecnico di Torino & Università degli Studi di Torino

BLOCKCHAIN TEE MACHINE LEARNING AUDITABILITY

Politecnico
di Torino

Noemi Romani
PhD student

The Core Challenge

Transparency Meets Confidentiality

ML algorithms increasingly support automated decisions while running off-chain on proprietary infrastructures with sensitive data and confidential models.

The goal: verifiable, auditable computations without disclosing confidential data or models.

Why It Matters

Enabling Adoption in Sensitive Domains

Healthcare and cyber-physical systems require verifiable ML workflows without exposing patient data, operational details, or sensitive infrastructure.

Minimizing ML Specific Risks

- Data poisoning
- Adversarial manipulation
- Model integrity

Building Trust in Decision Making Processes

Auditable execution supports regulatory compliance and strengthens stakeholder confidence in automated decisions.

The Technical Gap

Limitations of Individual Approaches

Blockchain

Immutable and auditable records;
but no guarantee on the correctness of off-chain computation.

Trusted Execution Environment (TEE)

Confidential execution and integrity,
but no cryptographic input-output binding.

Federated Learning (FL)

Protects training data privacy,
but remains exposed to poisoning and inference attacks.

Zero-Knowledge Proofs (ZKP)

Proves correctness without revealing data,
but carries high computational overhead.

TEE vs ZKP

Parameter	TEE	ZKP
Trust model	Hardware manufacturer	Mathematical / cryptographic
Execution overhead	1.1x-2x (up to 3x for enclave transitions)	100x-1.000.000x
Proof generation time	~ms	Seconds to hours
Proof size	1-2 KB	KB-MB
Input-output binding	No (attestation only)	Yes

The fundamental tension: ZKPs eliminate trust assumptions at the cost of computational scalability; TEEs achieve efficiency at the cost of relying on hardware trust.
Neither approach alone satisfies both requirements for complex ML pipelines.

Our Direction

Rather than treating these technologies as mutually exclusive, we propose integrating them strategically — each playing a *complementary* role in a unified architecture for auditable, privacy-preserving ML.

Three Complementary Building Blocks

TEE

Trusted Execution Environment

Certifies the execution of the ML algorithm, proving *model integrity* while preserving the *privacy* of both the model and the data.

Blockchain

Immutable on-chain record

Stores cryptographic hashes for the relevant ML pipeline components, enabling *independent verifiability* of the data stored on-chain as the effective result of the ML algorithm.

ZKPs

Zero-Knowledge Proofs

Binds the decisions taken on top of the ML output with the result stored on-chain, ensuring the *integrity* of the entire decision-making process.

State of the Art

How Existing Hybrids Combine Technologies

Blockchain + TEE

Strength

- Efficient confidential execution.

Limitation

- No cryptographic binding between model output and decision.

Applications

IoT authentication, vaccine tracking, high-performance rollups.

Blockchain + ZKP

Strength

- Strong mathematical guarantees.

Limitation

- High computational cost for ML workloads.

Applications

Energy trading, IoT privacy, federated learning verification.

Blockchain + TEE + ZKP

Strength

- Combines confidentiality and verifiability.

Limitation

- Mostly focused on execution verification.

Applications

Medical AI, generative AI inference, healthcare data governance.

Our Contribution:
From Execution to Decision Verifiability

1 Decision-focused verification

ZKPs certify invariants related to the decision-making process, not the full computation of a ML algorithm, such as:

- consistency between the computed output and the value registered on-chain;
- correctness of threshold-based decisions;
- compliance with predefined decision rules.

2 Strategic separation of concerns

Blockchain, TEE, and ZKPs act as complementary layers, each with a defined role, supported by decentralized IPFS storage.

3 Decision-level verifiability on top of TEE attestation

TEE attestation proves a binary ran in an enclave.

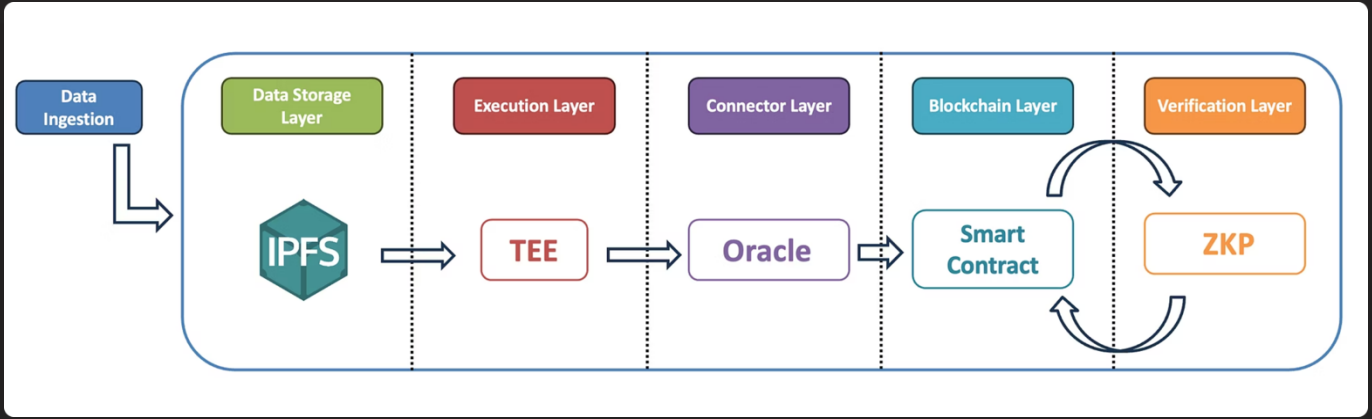
The ZKP layer adds a mathematical proof that the decision adhered to intended rules.

Blockchain, by storing cryptographic hashes associated to the relevant components of the ML pipeline, provides the opportunity for independent verifiability of data stored on chain as effective result of the ML algorithm.

System Architecture: A Hybrid Design

The architecture separates execution, verification, and auditability — assigning each technology a specific role.

- IPFS stores data via content identifiers;
    - TEE runs the model confidentially;
        - The Oracle relays hashes of significant data (computation output, CIDs of datasets and attestation report) on chain;
            - The Smart contracts store cryptographic hashes and verify proofs;
                - ZKPs verify selected decision properties;



Data Flow Across the Framework

01

Data Ingestion

Input *D* stored in IPFS, generating a unique CID.

02

Model Execution

Model *M* runs in the TEE producing output *O* with a remote attestation report.

03

Output Registration

Hash of *O*, CIDs related to attestation and data references are submitted on chain.

04

Decision Derivation

Function *f(O)* yields the final decision or alert.

05

ZKP Generation & Verification

Proof certifies the decision satisfies conditions; smart contracts verify it on-chain.

What is actually proven?

Hash(O) = H
and
Decision = f(O)

without revealing O.

Conclusion & Future Work

What We Established

- A hybrid architecture separating execution, verification, and auditability
- Decision-level verification as a practical alternative to full computation proofs
- A scalable trade-off between efficiency, confidentiality, and verifiability

Looking Ahead

- Prototype implementation and experimental evaluation
- Integration of efficient ZKP schemes for decision verification
- Assessment of performance and scalability across ML models and domains

This conceptual framework offers a scalable, practical path to auditable machine learning in real-world, privacy-sensitive applications.

Thank you for your attention!

Politecnico di Torino

Noemi Romani

PhD student