Makerere University

College of Computing and Information Sciences

Department of Computer Science

A Blockchain-Secured Federated Learning System for Detection and Early Warning of Foot-and-Mouth Disease in Ugandan Cattle Farms

Research Proposal

By

Muhindo Mubaraka

Student Number: 2400725633

Registration Number: 2024/HD05/25633U

muhindo.mubaraka@students.mak.ac.ug

Supervisor

Dr. Chongomweru Halimu

Department of Computer Science

A Research Proposal Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Computer Science

November 29, 2025

Abstract

Training machine learning models traditionally requires collecting all data in one central location. While this works in well-resourced environments with reliable internet, it creates problems in resource-limited settings. Privacy concerns prevent farmers and health facilities from sharing sensitive data, and poor connectivity makes uploading large datasets impractical. Centralizing data also creates security risks and trust issues when multiple organizations need to collaborate.

Federated learning offers an alternative by training models on distributed data without moving it from its source. Each participant trains locally and shares only model updates, which are aggregated to create a global model. However, this introduces a trust problem: participants cannot verify whether aggregation was done correctly or whether some contributed manipulated updates.

This research addresses the trust problem by integrating blockchain technology with federated learning. The blockchain provides a transparent, tamper-proof record of the training process, allowing participants to verify that all steps were executed correctly. I design and evaluate this system specifically for resource-limited environments, using Foot-and-Mouth Disease prediction in Ugandan cattle farms as a practical test case.

The system demonstrates that collaborative machine learning can work in settings with limited resources and unreliable connectivity while maintaining privacy and trust. The architecture balances security with computational efficiency, making blockchain-verified federated learning practical for developing regions. This has broader implications for privacy-preserving AI in agriculture, healthcare, and other domains where data sharing faces technical and trust barriers.

Keywords: Federated Learning, Blockchain, Privacy-Preserving Machine Learning, Disease Prediction, Distributed Systems, Differential Privacy, Byzantine Fault Tolerance

Contents

CHAPTER 1: INTRODUCTION **4**
0.1 Background of the Study . . . 4
0.2 Problem Statement . . . 4
0.3 Research Objectives . . . 5
0.3.1 Main Objective . . . 5
0.3.2 Specific Objectives . . . 5
0.4 Research Questions . . . 5
0.5 Significance of the Study . . . 6
0.6 Scope and Limitations . . . 6
0.6.1 Scope . . . 6
0.6.2 Limitations . . . 7

CHAPTER 2: LITERATURE REVIEW **8**
0.7 Introduction . . . 8
0.8 Federated Learning and Privacy Preservation . . . 8
0.9 Trust and Verification Challenges . . . 8
0.10 Blockchain Integration for Verifiable Federated Learning . . . 9
0.11 Applications in Disease Prediction and Healthcare . . . 9
0.12 Resource-Constrained Environments and Computational Efficiency . . . 10
0.13 Summary of Research Gaps . . . 10
0.14 Chapter Summary . . . 11

CHAPTER 3: RESEARCH METHODOLOGY **12**
0.15 Introduction . . . 12
0.16 Research Design . . . 12
0.17 System Architecture . . . 12
0.18 Prototype Implementation . . . 13
0.19 Performance Evaluation . . . 13

1 Timeline and Work Plan **13**
1.1 Phase 1: Literature Review and Data Preparation (Months 1-2) . . . 13
1.2 Phase 2: Core System Implementation (Months 3-5) . . . 14
1.3 Phase 3: Security and Privacy Implementation (Months 6-7) . . . 14
1.4 Phase 4: Adaptive Mechanisms (Month 8) . . . 14
1.5 Phase 5: Real-World Testing and Dashboard (Month 9) . . . 14
1.6 Phase 6: Comprehensive Evaluation (Months 10-11) . . . 15
1.7 Phase 7: Documentation and Dissemination (Month 12) . . . 15

1.8 Expected Contributions and Deliverables 15
1.8.1 Scientific Contributions 15
1.8.2 Technical Contributions 15
1.8.3 Practical Impact for Uganda 16
1.8.4 Global Relevance 16
1.9 Evaluation Metrics 16
1.9.1 Model Performance Metrics 16
1.9.2 Privacy Metrics 16
1.9.3 Fairness Metrics 16
1.9.4 Efficiency Metrics 17
1.9.5 Robustness Metrics 17

2 Ethical Considerations **17**
2.1 Data Privacy and Protection 17
2.2 Informed Consent and Participation 17
2.3 Responsible Use of Predictions 17
2.4 Equitable Access 18
2.5 Environmental Impact 18

3 Research Scope **18**
3.1 Acknowledged Limitations 18
3.2 Future Research Directions 19

4 Process Workflow **19**

5 Conclusion **19**

CHAPTER 1: INTRODUCTION

0.1 Background of the Study

Traditional machine learning requires collecting all data in one central location for training. In many real-world situations, especially in low-resource regions, this approach is not practical. There are several reasons why this does not work well.

First, privacy concerns make people reluctant to share their data. For example, a livestock farmer may fear to share photos or health records of their animals because they worry that government auditors or the ministry of agriculture will identify their farm and penalize them for any violations or diseases found. Similarly, hospitals are hesitant to share patient data because of confidentiality laws and fear of exposing sensitive medical information that could be traced back to individual patients or institutions.

Second, many organizations do not have the network and computing resources to send large amounts of data to a central server. For example, a rural health clinic with slow internet connection cannot afford to upload thousands of high-resolution medical images every day. The upload would take too long, consume too much bandwidth, and may fail frequently due to unreliable connectivity. This makes centralized machine learning not practical for such settings [2].

Federated learning (FL) solves these problems by keeping data on local devices and only sharing model updates instead of raw data. Each participant trains a model on their own data, then sends only the learned patterns (model updates) to a central server. The server combines these updates to create a global model without ever accessing the actual data [1, 3].

However, FL has a critical trust issue. When participants send their model updates, there is no easy way to verify if someone is sending corrupted or fake updates that could damage the model. This research looks at how to add a verification system using blockchain to create reliable records of who contributed what, while keeping the system light enough to work in places with limited resources.

0.2 Problem Statement

Federated learning has a trust problem. When multiple participants train a model together, there is no reliable way to verify if someone is sending honest updates or corrupted data. A malicious participant could poison the entire model by sending fake updates. Similarly, participants cannot verify if the central server is actually using their contributions or ignoring them.

Federated learning was designed to work in low-resource environments where participants have limited computing power and unreliable network connections. However,

there is no lightweight verification mechanism that can provide trust and accountability in these resource-limited settings.

The core problem is: how can I design a verification mechanism that provides trust and accountability in federated learning, while remaining lightweight enough to work in resource-constrained environments? This research addresses this gap by designing a practical blockchain-based verification system for federated learning in resource-limited settings.

0.3 Research Objectives

0.3.1 Main Objective

To design and evaluate a blockchain-based federated learning system with early warning capabilities for Foot-and-Mouth Disease in Ugandan cattle farms.

0.3.2 Specific Objectives

1. To design a system architecture that integrates federated learning with blockchain for secure and verifiable model aggregation.

2. To implement a lightweight prototype that works efficiently in resource-limited environments typical of Ugandan cattle farms.

3. To evaluate the system performance in terms of prediction accuracy, computational efficiency, and security against malicious attacks.

4. To assess the system's practical effectiveness through field testing with real Foot-and-Mouth Disease data from Ugandan cattle farms.

0.4 Research Questions

1. How can blockchain technology be integrated with federated learning to provide secure and verifiable model aggregation for early warning of Foot-and-Mouth Disease?

2. What system architecture will work efficiently in resource-limited environments typical of Ugandan cattle farms while maintaining prediction accuracy?

3. How does the system perform in terms of FMD prediction accuracy, computational cost, and security against malicious participants?

4. What is the practical effectiveness of the system when tested with real Foot-and-Mouth Disease data from Ugandan cattle farms?

0.5 Significance of the Study

This research addresses the trust problem in federated learning by combining it with blockchain technology to create a verifiable and secure system for collaborative machine learning in resource-limited environments.

The significance of this study extends across multiple dimensions:

Scientific Contribution: This research provides a practical solution to the trust problem in federated learning by integrating blockchain verification in a way that works efficiently in resource-limited settings. It shows how to balance security with computational efficiency.

Technical Innovation: The system demonstrates that blockchain-based federated learning can work in environments with limited resources and unreliable connectivity. This opens possibilities for similar applications in other low-resource settings.

Practical Application: By addressing Foot-and-Mouth Disease prediction in Ugandan cattle farms, this work solves a real problem in the livestock sector. Early warning can help farmers and veterinary officers respond quickly to disease outbreaks, reducing economic losses.

Data Privacy and Trust: The system allows multiple districts to collaborate on disease prediction without sharing raw data. The blockchain layer provides proof that the learning process was not manipulated, addressing the trust issue that prevents data collaboration.

Broader Applicability: While tested with FMD data in Uganda, the architecture can be adapted to other diseases and other countries facing similar challenges with distributed data and limited resources.

0.6 Scope and Limitations

0.6.1 Scope

This research focuses on:

- **Geographic Focus:** Uganda, specifically districts that participate in the Uganda Livestock Information Tracking System (ULITS) which covers FMD data collection
- **Disease Focus:** Foot-and-Mouth Disease as the test case for the early warning system
- **Data Period:** Historical FMD outbreak data from 2011-2024 for training and testing the system
- **Technical Focus:** Design and evaluation of a system that integrates federated learning with blockchain verification, specifically for resource-limited environments

0.6.2 Limitations

- **Data Quality:** The accuracy of predictions depends on the quality and completeness of FMD data collected by veterinary officers in different districts. Incomplete or inaccurate data will affect the system's performance.
- **Network Requirements:** While designed for low-connectivity environments, the system still needs periodic internet access for participants to share model updates and synchronize with the blockchain.
- **Scale of Testing:** The evaluation will involve up to 50 participating nodes (representing districts). Testing at much larger scales is beyond the scope of this research.
- **Trust Assumption:** The security mechanisms assume that most participants (at least 51%) act honestly. If the majority of participants are malicious, the system's security guarantees may not hold.
- **Single Disease:** While the system architecture can potentially work for other diseases, this research focuses only on Foot-and-Mouth Disease. Testing with other diseases is not included.

CHAPTER 2: LITERATURE REVIEW

0.7 Introduction

This chapter examines existing research on federated learning, blockchain technology, and their use in disease prediction. The review is organized around key themes: privacy-preserving machine learning, trust verification in distributed systems, blockchain integration, and practical applications in resource-limited settings. Through this analysis, I identify gaps that motivate this research.

0.8 Federated Learning and Privacy Preservation

Federated learning has gained attention as a way to train machine learning models without collecting data in one central location [1]. The basic idea is simple: instead of sending data to a server, each participant trains a model locally and only shares the model updates. The server aggregates these updates to improve the global model [2].

This approach addresses privacy concerns because raw data never leaves its source. For example, hospitals can collaborate on disease prediction without sharing patient records, and farms can work together on pest detection without exposing their operational data. This is particularly valuable in sectors where data is sensitive or regulations prevent data sharing.

However, while federated learning protects privacy during training, it does not solve the trust problem. Participants cannot verify whether the aggregation was done correctly or whether some participants sent manipulated updates. The system relies on trusting the central server, which becomes a single point of failure. Several researchers have pointed out this limitation [3], noting that federated learning alone does not provide proof that the process was fair or that the final model was not compromised.

0.9 Trust and Verification Challenges

The trust problem in federated learning becomes critical in real-world applications. When multiple organizations collaborate on model training, they need assurance that the process is transparent and verifiable. Without this, participants may hesitate to join the system, especially when decisions based on the model have significant consequences.

Recent work has explored different approaches to address trust. Some researchers propose cryptographic methods like secure multi-party computation and homomorphic encryption to verify computations without revealing data. Others suggest reputation systems where participants build trust over time based on their behavior. However, these approaches either add significant computational overhead or take time to establish trust, neither of which works well in resource-limited environments that need quick deployment.

The agricultural sector in developing countries presents a particular challenge. Veterinary officers collecting disease data in different districts may not trust a centralized system controlled by one authority. They need transparency about how their data contributes to predictions and confidence that their contributions are properly acknowledged. This trust gap prevents many potential collaborative systems from being adopted.

0.10 Blockchain Integration for Verifiable Federated Learning

Blockchain technology has emerged as a potential solution to the trust problem in federated learning. The idea is to use blockchain as a transparent, tamper-proof record of the training process. Each model update, aggregation step, and participant contribution can be recorded on the blockchain, creating an auditable history.

Several recent studies have explored this integration. Zhang et al. [4] proposed combining blockchain with federated learning using cryptographic methods like verifiable random functions and zero-knowledge proofs. Their system aims to ensure fairness and privacy during model training. However, their work focuses on the theoretical framework and does not address practical deployment in resource-limited environments.

Chen et al. [5] developed a system called FLock that uses blockchain state channels to prevent model poisoning attacks. Their approach shows that blockchain can help detect malicious participants who try to corrupt the model. The system is robust against attacks, but it is designed for environments with good network connectivity and sufficient computational resources.

What I observe from these studies is that blockchain can indeed provide transparency and verification for federated learning. The challenge is making it work efficiently in settings where resources are limited. Most proposed systems assume stable internet connections and powerful devices, which does not match the reality in many developing regions.

0.11 Applications in Disease Prediction and Healthcare

Federated learning has been applied to healthcare and disease prediction with promising results. Teo et al. [6] reviewed 612 studies on federated learning in healthcare and found that it enables hospitals to collaborate on diagnostic models without violating patient privacy. The review highlights successes in medical image analysis, disease outbreak prediction, and drug discovery.

However, most healthcare applications focus on well-resourced medical institutions in developed countries. The infrastructure requirements, including high-speed networks and powerful computing devices, are often not available in rural clinics or agricultural settings in developing regions.

In the livestock sector, Kapalaga et al. [7] created a unified dataset for Foot-and-Mouth Disease prediction in Uganda. Their work addresses data fragmentation by combining records from multiple sources into one national dataset. They show that machine learning models can predict FMD outbreaks based on factors like rainfall and temperature. However, their approach is centralized, requiring all data to be collected in one location. This creates privacy concerns and network bottlenecks, especially in areas with poor connectivity.

What is missing from the literature is a system that combines the privacy benefits of federated learning with the trust guarantees of blockchain, specifically designed for disease prediction in resource-limited agricultural settings. The existing work either focuses on healthcare in developed countries or uses centralized approaches that do not address privacy and trust issues.

0.12 Resource-Constrained Environments and Computational Efficiency

An important consideration for any system deployed in developing regions is computational efficiency. Many proposed blockchain-federated learning systems use complex cryptographic protocols that require significant processing power and memory. This works in data centers but not on the mobile devices or low-power computers typically available in rural areas.

Some researchers have started exploring lightweight approaches. Edge computing techniques show promise for running machine learning models on resource-limited devices. Model compression and quantization can reduce the size and computational requirements of neural networks. However, most of this work focuses on inference (using trained models) rather than training, and little attention has been paid to making the blockchain verification layer lightweight.

The gap here is clear: a system architecture is needed that balances security and verifiability with computational efficiency. The system must work with intermittent internet connectivity, limited processing power, and minimal storage capacity while still providing the transparency and trust benefits of blockchain.

0.13 Summary of Research Gaps

Through this literature review, I identified three critical gaps that this research addresses:

1. **Trust Verification Gap:** While federated learning provides privacy, existing systems lack built-in mechanisms for participants to verify that model aggregation was done correctly and that all participants followed the protocol. Current solutions ei-

ther rely on trust in a central authority or use cryptographic methods too heavy for resource-limited settings.

2. **Practical Deployment Gap:** Most proposed blockchain-federated learning systems are designed for well-resourced environments. There is no system specifically designed for resource-limited agricultural settings with poor connectivity, limited devices, and the need for disease early warning.

3. **Application Gap:** Existing applications focus on healthcare in developed countries or use centralized approaches in agriculture. No system combines blockchain-verified federated learning specifically for livestock disease prediction in developing regions where data is distributed, trust is a concern, and resources are limited.

0.14 Chapter Summary

This review shows that both federated learning and blockchain technologies have advanced significantly, but their integration for practical early warning systems in resource-constrained environments remains unexplored. The trust problem in federated learning is well recognized but not adequately solved for settings like Ugandan cattle farms. The next chapter describes how I design and evaluate a system that addresses these gaps.

CHAPTER 3: RESEARCH METHODOLOGY

0.15 Introduction

This chapter describes the methodology for designing, implementing, and evaluating a blockchain-secured federated learning system for Foot-and-Mouth Disease early warning in Uganda. The approach follows a systematic process: architecture design, prototype implementation, and comprehensive evaluation across accuracy, efficiency, and security dimensions.

0.16 Research Design

This study adopts a design science research methodology, which is appropriate for creating and evaluating new artifacts in information systems. The research follows an iterative process of:

1. Problem identification and motivation
2. Objectives definition
3. Design and implementation
4. Demonstration
5. Evaluation
6. Communication

0.17 System Architecture

I will design a layered architecture that strategically separates concerns while maintaining integration between federated learning and blockchain components:

- **Data Layer:** Local data remains on farm systems with standardized interfaces for secure access
- **Federated Learning Layer:** Model training occurs locally with secure aggregation protocols
- **Blockchain Layer:** Hyperledger Fabric will provide the trust foundation for recording model hashes and verification
- **Application Layer:** Early warning dashboard and alert system

0.18 Prototype Implementation

The system will be implemented using technologies suitable for resource-limited environments:

- **PySyft** for federated learning, chosen for its support of privacy-preserving operations and secure aggregation
- **Hyperledger Fabric** for blockchain verification, selected for its practical Byzantine fault tolerance and lower computational requirements compared to proof-of-work systems
- **Django** for backend services and **React.js** for the web-based early warning dashboard

0.19 Performance Evaluation

I will conduct comprehensive testing across three key dimensions:

1. **Model Accuracy:** Compare prediction performance using F1-score, precision, and recall metrics
2. **System Efficiency:** Measure communication overhead, training convergence time, and computational resource usage
3. **Security Analysis:** Test robustness against model poisoning attacks and privacy preservation

1 Timeline and Work Plan

The research will span 12 months, with each phase building on previous work to ensure systematic progress.

1.1 Phase 1: Literature Review and Data Preparation (Months 1-2)

The first two months focus on establishing a solid foundation. This includes comprehensive literature review, finalizing the system architecture design, and obtaining access to the FMD dataset from Uganda Livestock Information Tracking System (ULITS).

Data preparation involves cleaning, standardizing formats, and creating the unified dataset that will be distributed across simulated nodes for federated training experiments.

1.2 Phase 2: Core System Implementation (Months 3-5)

This phase involves building the federated learning infrastructure. The approach is incremental: first implementing a basic federated learning system to ensure correct model training and aggregation, then integrating the blockchain layer for verification.

Building components separately before integration reduces complexity and makes debugging more manageable. By month 5, both the federated learning and blockchain components should be functional and integrated.

1.3 Phase 3: Security and Privacy Implementation (Months 6-7)

This phase adds privacy preservation and attack resistance mechanisms. Differential privacy will be implemented and calibrated to balance privacy protection with model accuracy.

For differential privacy, the key equation is:

$$\Pr[\mathcal{M}(D) \in S] \leq e^{\epsilon} \cdot \Pr[\mathcal{M}(D') \in S] + \delta \tag{1}$$

where $\mathcal{M}$ is the mechanism, D and D' are neighboring datasets differing by one record, ϵ controls privacy loss, and δ is a small failure probability.

For aggregation, if there are N clients and client i sends update θ_i with local dataset size n_i, standard federated averaging computes:

$$\theta_{global} = \frac{\sum_{i=1}^{N} n_i \theta_i}{\sum_{i=1}^{N} n_i} \tag{2}$$

1.4 Phase 4: Adaptive Mechanisms (Month 8)

This phase implements mechanisms for detecting and handling distribution shifts. For each district d at time t, the system tracks validation loss $L_d^{(t)}$ over a sliding window. If the moving average increases beyond a threshold:

$$\frac{1}{w} \sum_{k=t-w}^{t} L_d^{(k)} > \mu + 2\sigma \tag{3}$$

where μ and σ are historical mean and standard deviation, this signals distribution drift requiring model retraining.

1.5 Phase 5: Real-World Testing and Dashboard (Month 9)

This phase tests the system under realistic conditions including slow internet connectivity, node failures, and participants dropping offline during training. The early warning

dashboard will be developed to allow veterinary officers to view predictions with explanations.

Testing focuses on robustness under the actual conditions expected in Ugandan districts, ensuring the system works despite infrastructure limitations.

1.6 Phase 6: Comprehensive Evaluation (Months 10-11)

This phase conducts systematic experiments across all evaluation dimensions: prediction accuracy compared to centralized training, system efficiency under varying network conditions, and security robustness against different attack scenarios (including 30-40% malicious participants).

Performance metrics will be collected and analyzed to document the trade-offs between privacy, security, and efficiency.

1.7 Phase 7: Documentation and Dissemination (Month 12)

The final month focuses on thesis writing, preparing research papers for publication, and releasing the open-source code and documentation. All datasets and implementation details will be made publicly available to enable replication and extension of this work.

1.8 Expected Contributions and Deliverables

1.8.1 Scientific Contributions

This research provides a practical solution to the trust problem in federated learning by integrating blockchain verification in resource-limited settings. The system architecture, implementation strategies, and benchmark performance data will serve as a foundation for future work in privacy-preserving distributed machine learning.

The research demonstrates how to balance security, privacy, and computational efficiency in environments with limited resources and unreliable connectivity.

1.8.2 Technical Contributions

The open-source implementation will include production-ready code for blockchain-verified federated learning, complete documentation, and deployment guidelines. Researchers and practitioners can adapt this system to their own applications without building from scratch.

All components will be modular to allow customization for different use cases, datasets, and infrastructure constraints.

1.8.3 Practical Impact for Uganda

The system provides a working early warning mechanism for Foot-and-Mouth Disease that could help veterinary officers respond quickly to outbreaks. Beyond the specific application, it demonstrates that Uganda can benefit from advanced AI without centralizing sensitive data or relying on external infrastructure.

This establishes a model for how developing countries can maintain data sovereignty while leveraging collaborative machine learning.

1.8.4 Global Relevance

The challenges addressed—intermittent connectivity, limited computational resources, privacy requirements, and trust verification—are common across developing regions. Success in the Ugandan context provides evidence that sophisticated machine learning systems can work in similar conditions worldwide.

The approach applies beyond agriculture to healthcare, disaster response, and other domains where distributed data and limited resources create barriers to AI adoption.

1.9 Evaluation Metrics

1.9.1 Model Performance Metrics

Precision and Recall: For binary classification:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

F1-Score: The harmonic mean:

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

Target: F1-score above 0.80, AUC-ROC above 0.85.

1.9.2 Privacy Metrics

Differential Privacy Budget: Track cumulative ϵ over training rounds, targeting $\epsilon < 10$.

Membership Inference Attack Success Rate: Measure if attackers can identify whether specific farm data was in training set.

1.9.3 Fairness Metrics

Gini Coefficient: Measures inequality in district selection frequency. Target below 0.3.

Performance Equity: Variance in F1-scores across districts. Lower variance means more equitable performance.

1.9.4 Efficiency Metrics

Communication Overhead Ratio:

$$\text{Overhead Ratio} = \frac{\text{Bytes in FL}}{\text{Bytes in Centralized}} \tag{6}$$

Expect ratio less than 0.1 (90% bandwidth reduction).

Training Time: Target under 30 minutes per round on realistic hardware.

1.9.5 Robustness Metrics

Accuracy Under Attack: For p fraction of malicious clients:

$$\text{Robustness Score} = \frac{\text{Accuracy with attack}}{\text{Accuracy without attack}} \tag{7}$$

Aim to maintain above 0.80 even with 40% attackers.

2 Ethical Considerations

2.1 Data Privacy and Protection

All personally identifiable information will be removed from datasets before use. The federated learning approach ensures that raw farm-level data never leaves its source location. Limitations of privacy guarantees will be clearly documented and communicated to participants.

2.2 Informed Consent and Participation

If human subjects are involved, university ethics board approval will be obtained before data collection. Participants will receive clear explanations of how the system works, what data is used, potential risks, and their right to withdraw at any time without penalty.

2.3 Responsible Use of Predictions

The system provides decision-support tools, not autonomous decisions. Veterinary officers maintain authority and responsibility for final judgments. Predictions include explanations to help users evaluate reliability and make informed decisions.

2.4 Equitable Access

All software and research outputs will be released under open-source licenses to ensure broad access. The system is designed to work with intermittent connectivity so that all districts, regardless of infrastructure quality, can benefit equally.

2.5 Environmental Impact

Hyperledger Fabric uses practical Byzantine fault tolerance rather than energy-intensive proof-of-work consensus. Federated learning reduces network traffic and data center energy consumption by processing data locally.

3 Research Scope

This research focuses on designing and evaluating a blockchain-verified federated learning system for FMD prediction in Uganda. The scope includes:

- **Geographic Coverage:** Ugandan districts participating in the Uganda Livestock Information Tracking System (ULITS)
- **Disease Application:** Foot-and-Mouth Disease early warning as the primary test case
- **Data Timeline:** Historical FMD outbreak data from 2011-2024 for training and evaluation
- **Technical Components:** Federated learning with blockchain verification, differential privacy, secure aggregation, and adaptation to distribution shifts

3.1 Acknowledged Limitations

Production Deployment: Full operational deployment across Uganda requires additional resources, partnerships, and infrastructure beyond this research scope.

Security Testing: Evaluation covers common Byzantine attack scenarios but cannot exhaustively test all possible threat vectors.

Trust Model: The system assumes an honest majority (less than 50% malicious participants), which is a fundamental requirement for Byzantine fault-tolerant consensus.

Scale: Testing involves up to 50 simulated nodes representing districts. Performance at larger scales may differ.

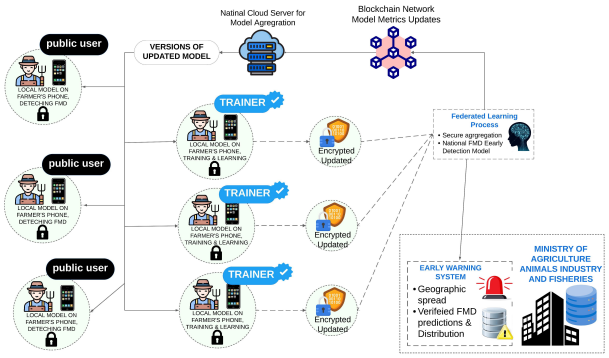


Figure 1: System workflow showing local data preparation, federated model training, blockchain verification, secure aggregation, and early warning generation. Data never leaves local farms while maintaining verifiable trust through blockchain recording.

3.2 Future Research Directions

This foundation enables future work on multi-disease monitoring, cross-country collaboration, real-time sensor integration, personalized district-level models, and economic incentive mechanisms for sustained participation.

4 Process Workflow

The workflow in Fig. 1 ensures data privacy while maintaining accountability through blockchain-recorded model updates and aggregation steps.

5 Conclusion

This research addresses the trust problem in federated learning by integrating blockchain verification, specifically designed for resource-limited environments. The system enables collaborative machine learning without centralizing data, making it practical for settings where privacy concerns, limited connectivity, and infrastructure constraints prevent traditional approaches.

Using Foot-and-Mouth Disease prediction in Ugandan cattle farms as a test case provides real-world validation under challenging conditions. The FMD application addresses a genuine need for early warning in the livestock sector while demonstrating that blockchain-verified federated learning can work despite intermittent internet, limited computing resources, and distributed data sources.

The research contributes a working system architecture, implementation strategies, and empirical evidence for privacy-preserving collaborative machine learning in developing regions. Beyond the specific application to FMD in Uganda, the approach applies to

healthcare, agriculture, and other domains where data remains distributed due to privacy, sovereignty, or infrastructure limitations.

The outcome includes open-source tools, performance benchmarks, and practical lessons that advance the field of privacy-preserving machine learning while addressing a real problem in Uganda's livestock sector.

References

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proc. 20th Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, 2017.

[2] P. Kairouz *et al.*, "Advances and Open Problems in Federated Learning," *Foundations and Trends in Machine Learning*, vol. 14, no. 1–2, pp. 1–210, 2021.

[3] K. Bonawitz *et al.*, "Practical Secure Aggregation for Privacy-Preserving Machine Learning," in *Proc. ACM SIGSAC Conf. Computer and Communications Security*, 2017.

[4] Y. Liu, X. Qu, and G. Chen, "Blockchain-Based Practical and Privacy-Preserving Federated Learning with Verifiable Fairness," *Mathematics*, vol. 11, no. 5, p. 1091, 2023.

[5] R. Chen, Y. Li, and M. Zhang, "FLock: Robust and Privacy-Preserving Federated Learning based on Practical Blockchain State Channels," *Cryptology ePrint Archive*, Paper 2024/1797, 2024.

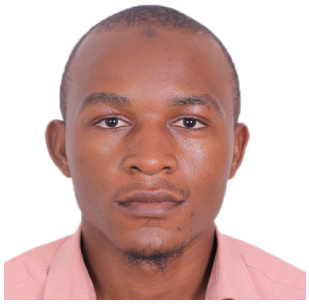
[6] R. S. Antunes, C. A. da Costa, and A. Küdde, "Federated Machine Learning in Healthcare: A Systematic Review," *ACM Computing Surveys*, vol. 55, no. 5, pp. 1–35, 2022.

[7] T. Kapalaga, M. Mubangizi, and P. Kisaakye, "A Unified Foot and Mouth Disease Dataset for Uganda: Evaluating Machine Learning Predictive Performance Degradation Under Varying Distributions," *Frontiers in Artificial Intelligence*, vol. 7, 2024.

[8] C. Dwork and A. Roth, "The Algorithmic Foundations of Differential Privacy," *Foundations and Trends in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[9] Hyperledger Foundation, "Hyperledger Fabric," [Online]. Available: https://www.hyperledger.org/use/fabric

About the Author



Muhindo Mubaraka received his B.Sc. in Computer Science from the Islamic University of Technology, Bangladesh, in 2022. He is currently pursuing a Master of Science in Computer Science at Makerere University. He works as a software engineer and his research interests include federated learning and blockchain technology.