International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 1, Jan-Feb 2026
Available at wwwijsred.com

RESEARCH ARTICLE OPEN ACCESS

Self-Adaptive AI Systems Using Federated Learning and Autonomous Model Governance

Albin SS
Department of Computer Science and Engineering Narayanaguru college of engineering
Email: ss.albin20@email.com

----------**********----------

Abstract :

The rapid adoption of Artificial Intelligence (AI) systems across distributed environments such as cloud platforms, edge devices, and Internet of Things (IoT) networks has raised serious concerns related to data privacy, security, fairness, and regulatory compliance. Traditional centralized machine learning approaches require the aggregation of sensitive data into a single location, which is often impractical due to legal, ethical, and security constraints. Federated Learning (FL) has emerged as a promising decentralized learning paradigm that enables collaborative model training without sharing raw data.

However, existing federated learning systems face several limitations, including model drift, bias accumulation, lack of adaptability to dynamic data distributions, and the absence of automated governance mechanisms. These challenges reduce long-term reliability and trustworthiness of AI systems deployed in real-world environments.

This paper proposes a self-adaptive AI framework that integrates federated learning with autonomous model governance. The proposed system continuously monitors performance metrics, fairness indicators, and compliance constraints, and automatically adapts training strategies without human intervention. Autonomous governance mechanisms are embedded directly into the learning lifecycle to ensure ethical, transparent, and compliant AI behavior. The proposed approach improves robustness, scalability, and trustworthiness, making it suitable for privacy-sensitive and distributed applications.

Keywords — Self-Adaptive AI, Federated Learning, Autonomous Model Governance, Trustworthy AI, Privacy-Preserving Learning

----------**********----------

I. INTRODUCTION

Artificial Intelligence (AI) systems are increasingly deployed in distributed and data-intensive environments such as healthcare networks, financial systems, smart cities, and edge computing infrastructures. These environments generate massive volumes of sensitive and heterogeneous data. Centralized machine learning approaches require transferring this data to a central server, which introduces privacy risks, high communication costs, and regulatory challenges.

Federated Learning (FL) addresses these limitations by enabling multiple clients to collaboratively train a global model while keeping their data locally. Although federated learning improves privacy preservation, it introduces new challenges such as statistical data heterogeneity, model drift, biased updates, and limited adaptability to changing environments. Moreover, most existing federated learning frameworks rely on manual monitoring and governance, which is inefficient for large-scale autonomous systems.

As AI systems become more autonomous and critical to decision-making, there is a growing need for self-adaptive mechanisms that allow systems to monitor their own performance and adjust behavior accordingly. In addition, AI governance must be integrated directly into the system lifecycle to ensure fairness, transparency, accountability, and regulatory compliance.

This research focuses on designing a self-adaptive AI system that combines federated learning with autonomous model governance. The proposed

framework aims to achieve continuous learning, automated adaptation, and trustworthy AI behavior in distributed environments.

II. BACKGROUND AND RELATED WORK

A. Federated Learning

Federated learning is a decentralized machine learning approach where multiple clients train a shared global model without exchanging raw data. Each client computes local model updates using its private dataset, and only the model parameters are transmitted to a central aggregation server. This approach significantly reduces privacy risks and data transfer overhead.

Despite its advantages, federated learning faces several challenges. Data heterogeneity across clients leads to slow convergence and unstable performance. Communication constraints increase training latency. Additionally, federated systems are vulnerable to biased or malicious updates, which can degrade global model quality.

B. Self-Adaptive AI Systems

Self-adaptive AI systems are capable of monitoring their internal state and external environment and modifying their behavior autonomously. These systems aim to maintain optimal performance under changing conditions such as evolving data distributions or system constraints. Adaptation techniques include dynamic learning rate adjustment, selective client participation, and automated retraining.

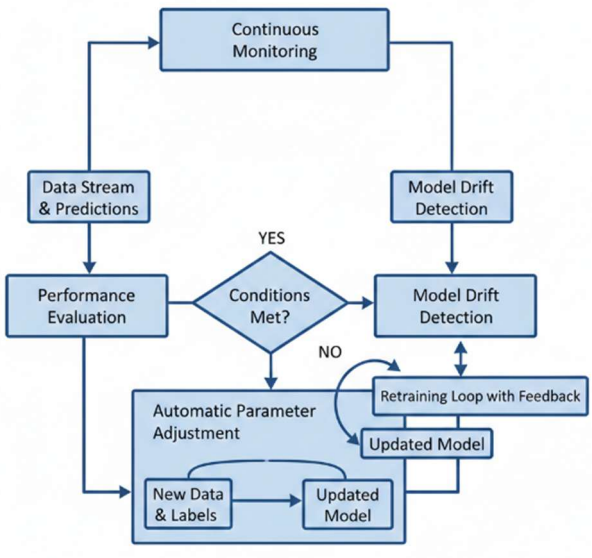

Fig. 1. Autonomous Model Governance Framework

C. AI Governance and Trustworthy AI

AI governance refers to frameworks and mechanisms that ensure ethical, fair, and transparent operation of AI systems. Trustworthy AI emphasizes principles such as fairness, accountability, explainability, and compliance with regulations. Existing governance mechanisms are largely manual and reactive, making them unsuitable for large-scale autonomous systems.

Component	Description
Client Nodes	Edge devices or organizational nodes that store local datasets and perform decentralized model training while preserving data privacy
Federated Learning Module	Decentralized learning mechanism that enables collaborative model training without sharing raw data
Self-Adaptive Learning Engine	Continuous monitoring unit that dynamically adjusts learning parameters based on system performance
Model Drift Detection Unit	Analytical module that identifies changes in data distribution and performance degradation
Autonomous Governance Module	Built-in governance mechanism responsible for fairness monitoring, bias control, and compliance enforcement
Secure Aggregation Server	Central coordination unit that aggregates encrypted model updates into a global model
Performance Analytics Dashboard	Visualization and reporting interface that presents accuracy, efficiency, and fairness metrics

III. PROBLEM STATEMENT AND OBJECTIVES

A. Problem Statement

Current federated learning systems lack autonomous mechanisms to handle model drift, bias accumulation, and regulatory compliance in dynamic environments. Manual governance and adaptation approaches are

insufficient to ensure long-term reliability, scalability, and trustworthiness of AI systems.

B. Objectives

The main objectives of this research are:

- To design a self-adaptive AI framework integrating federated learning and autonomous governance
- To enable continuous monitoring of model performance and fairness
- To implement automated adaptation strategies without human intervention
- To ensure ethical, transparent, and compliant AI behavior
- To evaluate robustness, scalability, and effectiveness of the proposed system

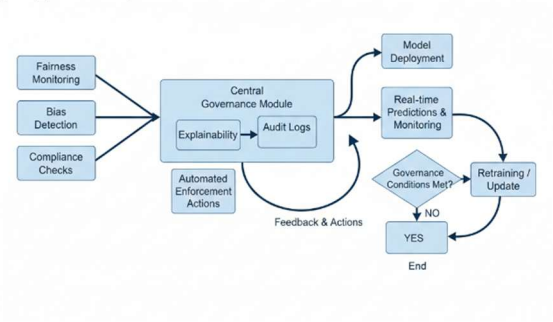

Fig. 2. Autonomous Model Governance Framework

IV. PROPOSED SYSTEM ARCHITECTURE

The proposed self-adaptive AI framework consists of four major components designed to work cohesively in distributed environments.

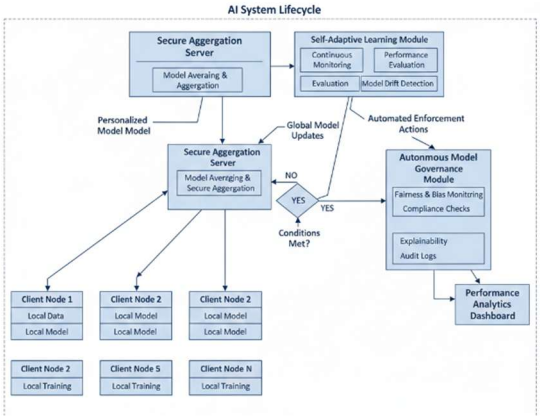

Fig. 3. System architecture

A. Client Layer

The client layer consists of edge devices, organizational nodes, or IoT devices that hold local datasets. Each client performs local model training using private data and applies governance constraints such as fairness checks before sending updates.

B. Federated Aggregation Server

The aggregation server collects encrypted model updates from clients and performs secure aggregation to generate a global model. The updated model is then redistributed to participating clients.

C. Self-Adaptive Learning Module

This module continuously monitors performance metrics such as accuracy, loss, and convergence rate. Based on observed trends, it automatically adjusts training parameters, learning rates, and client participation strategies.

D. Autonomous Model Governance Module

The governance module evaluates fairness, bias, and compliance constraints. If violations are detected, corrective actions such as retraining, update rejection, or rollback are triggered automatically. All decisions are logged to ensure transparency and auditability.

V. METHODOLOGY

A. Federated Learning Workflow

The workflow begins with global model initialization, followed by client selection and local training. Clients compute updates and securely transmit them to the aggregation server. The global model is updated using federated averaging and redistributed.

B. Adaptation Strategy

Performance metrics are continuously evaluated to detect concept drift or degradation. The system autonomously adapts training frequency, model parameters, and client weights to maintain optimal performance.

C. Governance Enforcement

Governance rules are applied at both client and server levels. Bias detection metrics and compliance policies ensure ethical behavior throughout the learning lifecycle.

VI. APPLICATIONS AND USE CASES

The proposed framework is suitable for various real-world applications.

In healthcare systems, it enables privacy-preserving learning from distributed medical data

while ensuring compliance with regulations. In financial services, it supports fraud detection and credit scoring with automated bias monitoring. In smart city environments, the framework enables ethical and scalable analytics across distributed sensors.

VII. CHALLENGES AND LIMITATIONS

The proposed framework introduces additional computational overhead and system complexity. Accurate monitoring metrics are critical for effective adaptation and governance. Scalability in very large federated networks remains a challenge.

VIII. FUTURE SCOPE

Future enhancements include integration with blockchain for transparent governance, reinforcement learning-based adaptation strategies, cross-domain federated governance, and real-world deployment studies.

IX. CONCLUSION

This paper presented a self-adaptive AI framework that integrates federated learning with autonomous model governance. By enabling continuous monitoring, automated adaptation, and built-in governance, the proposed system enhances trustworthiness, scalability, and sustainability of distributed AI systems. The framework represents a significant step toward responsible and autonomous AI in privacy-sensitive environments.

