NVIDIA FLARE

An open-source, domain-agnostic SDK for secure, privacy-preserving federated learning and distributed multi-party AI model collaboration.

https://nvidia.github.io/NVFlare/

Overview

NVIDIA FLARE (Federated Learning Application Runtime Environment) is a domain-agnostic, open-source, and extensible Python SDK designed to facilitate federated learning (FL) and federated analytics across diverse industries, particularly healthcare and finance.

Product Overview & Benefits FLARE enables researchers and data scientists to adapt existing machine learning (ML) and deep learning (DL) workflows to a federated paradigm with minimal code changes. The core value proposition is enabling collaborative AI model training across decentralized data sources without sharing the raw, sensitive data, thereby mitigating data security and privacy risks and ensuring regulatory compliance. It is built for robust, production-scale deployment, supporting scaling from a single-machine simulation to real-world, multi-site production environments in the cloud or on-premise.

Main Features & Capabilities

Privacy-Preserving Technologies (PETs): Includes differential privacy, homomorphic encryption, and private set intersection (PSI).
Model Agnostic: Supports all major ML/DL frameworks including PyTorch, TensorFlow, XGBoost, RAPIDS, Nemo, and NumPy.
FL Workflows: Built-in support for horizontal and vertical federated learning, as well as reference FL algorithms (e.g., FedAvg, FedProx, FedOpt, Scaffold, Ditto).
Developer Tools: Includes an FL Simulator for rapid prototyping, a Command Line Interface (CLI) for orchestration, and the FLARE Dashboard (a web-based UI) for simplified project management, deployment, and secure provisioning of client startup kits.
Security: Implements enterprise-grade security features like Mutual TLS (mTLS) authentication via Public Key Infrastructure (PKI), federated authorization, and built-in audit logs.

Extensibility: Features a fully customizable and modular architecture with an extensive API for developing new workflows and algorithms.
Target Users & Use Cases FLARE is primarily targeted at AI researchers, data scientists, and platform developers in large organizations. Key use cases include: Multi-Party AI Collaboration, Healthcare Research (e.g., medical imaging, genetic analysis), Financial Services (BFSI), and Autonomous Driving (Automotive).

Key Features

- Privacy-Preserving Technologies (Differential Privacy, HE, PSI)
- FL Simulator for Prototyping and Debugging
- FLARE Dashboard (Web UI for Project Management & Provisioning)
- Support for Horizontal and Vertical Federated Learning
- Extensible Python SDK with Modular Architecture
- Built-in Federated Learning Algorithms (FedAvg, FedProx, etc.)
- LLM Support via Streaming API
- Enterprise-grade Security (mTLS, Federated Authorization)

Pricing

Model: free

NVIDIA FLARE is an open-source SDK released under the Apache 2.0 license, available for free download via GitHub and PyPi. Enterprise-grade support is available through NVIDIA's commercial offerings (e.g., NVIDIA AI Enterprise).

Target Company Size: medium, enterprise

Integrations

PyTorch, TensorFlow, XGBoost, Nemo, RAPIDS, Numpy, MONAI

Compliance & Certifications

HIPAA-enabling, GDPR-enabling, X.509 Standard (for PKI credentials)