Privacy-Preserving Machine Learning with Federated Learning Training

Duration: **5 Day**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

Overview

Privacy-Preserving Machine Learning with Federated Learning Training is an advanced engineering and research-oriented program focused on developing **distributed machine learning models** in environments where collecting sensitive data in a central location is restricted or impractical.

Federated Learning (FL) enables multiple organizations or devices to collaboratively train a**machine learning model** without transferring their raw datasets to a central server. Instead, models are trained where the data resides, while **model updates** are shared to improve a common global model.

The course begins with the **FedAvg** algorithm and explores **Cross-Silo** and **Cross-Device** architectures, centralized and decentralized learning approaches, **Non-IID Data (Statistical Heterogeneity)**, and differences in computing and network resources.

Participants also examine methods such as **FedProx**, designed to address challenges associated with heterogeneous data, as well as **Privacy-Enhancing Technologies (PETs)** including **Differential Privacy (DP)**, **Homomorphic Encryption (HE)**, and **Secure Multi-Party Computation (SMPC)**.

Through hands-on activities with **Flower** and **TensorFlow Federated (TFF)**, participants gain practical exposure to distributed model training approaches that prioritize data privacy.

Prerequisites

There are **no mandatory prerequisites** for attending this training.

Who Should Attend

- **Data Privacy Engineers** and **Security Architects**
- **Infrastructure and MLOps Engineers**
- **Embedded Systems Specialists** and **Mobile Developers**
- **Senior AI and Machine Learning Engineers**
- **Machine Learning Researchers**
- Technical professionals working with sensitive data in healthcare, finance, and telecommunications
- **CTOs**, **Technology Leaders**, and **R&D Directors**

Privacy-Preserving Machine Learning with Federated Learning Training

Duration: **5 Day**

BİLGİNÇ IT Academy
info@bilginc.com
bilginc.com

What You Will Learn

Upon completing this training, you will be able to:

- Explain **Federated Learning** architectures and how they differ from centralized machine learning.
- Develop **collaborative machine learning models** across distributed datasets.
- Understand the **FedAvg** training cycle and the purpose of **FedProx**.
- Compare **Cross-Silo** and **Cross-Device** Federated Learning architectures.
- Evaluate challenges related to **Non-IID data** and **system heterogeneity**.
- Identify privacy-enhancing methods such as **Differential Privacy, Homomorphic Encryption, and SMPC**.
- Relate the data privacy requirements of **GDPR, HIPAA, and KVKK** to architectural decisions.
- Work with **Flower** and **TensorFlow Federated** to explore distributed model training.
- Design approaches that share **model updates rather than centralizing raw data**.
- Evaluate **privacy-preserving AI solutions** for industries that process sensitive information.

Outline

1. Foundations of Privacy-Preserving AI and Requirements Analysis

Limitations of Traditional Centralized Machine Learning

- Examining the **centralized machine learning** approach
- Understanding **data privacy risks**
  - Data breaches
  - Secondary use of data
  - Identity disclosure
- The impact of regulations such as **GDPR, HIPAA, and KVKK** on centralized data processing
- Challenges associated with **data silos**, collaboration, and competitive opportunities

Privacy-Enhancing Technologies (PETs)

- **Data anonymization** techniques and their limitations
- **Re-identification attacks**

BİLGİNÇ IT Academy

info@bilginc.com

bilginc.com

- **Differential Privacy (DP)**
- **Homomorphic Encryption (HE)**
- **Secure Multi-Party Computation (SMPC)**
- **Zero-Knowledge Proofs (ZKP)**
- The relationship between **Federated Learning** and other privacy-enhancing technologies

2. Federated Learning Architecture and Core Algorithms

How Federated Learning Works

- Fundamental components of **Federated Learning**
- Interaction between the central server and distributed **clients**
- Distributing models to participating clients
- Training models on local datasets
- Sharing **model weights and gradient updates**
- Aggregating updates to improve the **global model**

FedAvg Algorithm

- Understanding **Federated Averaging (FedAvg)**
- The local training and global aggregation cycle
- Collaborative model training without sharing raw data

Federated Learning Architectures

- **Cross-Silo Federated Learning:** Collaborative training across organizations
- **Cross-Device Federated Learning:** Training across large numbers of edge devices
- **Centralized Federated Learning**
- **Decentralized (Peer-to-Peer / P2P) Federated Learning**
- **Hierarchical Federated Learning**

3. Key Challenges in Federated Learning: Heterogeneity

- Understanding **Non-IID Data**
- **Statistical heterogeneity** across participating clients
- The effects of different data distributions on model training

BİLGİNÇ IT Academy

info@bilginc.com

bilginc.com

- **System heterogeneity**
  - Differences in computational resources
  - Network connectivity variations
  - Device and infrastructure limitations
- **Communication bottlenecks** and network costs
- The role of **FedProx** in heterogeneous learning environments

4. Enhancing Security: Differential Privacy and Cryptographic Methods

Differential Privacy

- Introduction to **Differential Privacy (DP)**
- Adding mathematical noise to model updates
- Reducing the risk of exposing individual data contributions

Cryptographic Privacy Techniques

- **Homomorphic Encryption (HE)** and computation on encrypted data
- Aggregating encrypted model updates
- **Secure Multi-Party Computation (SMPC)**
- Combining **Federated Learning** with privacy-enhancing technologies

5. Hands-On Workshop: Flower and TensorFlow Federated (TFF)

- Introduction to the **Flower** library
- Python-based, framework-agnostic **Federated Learning** approaches
- Exploring **TensorFlow Federated (TFF)**
- Practical evaluation of distributed model training
- Working with local model training and**global model updates**