ANESTIS DALGKITSIS ALEXANDROS KOUFAKIS

 University of Amsterdam

JORRIT STUTTERHEIM, ALEANDRO MIFSUD, PRIYANKA ATWANI, LEON GOMMANS, CEES DE LAAT, CHRYSA PAPAGIANNI, ANA OPRESCU

SECURE COLLABORATIVE MODEL TRAINING WITH VFL IN MULTI-DOMAIN ENVIRONMENTS

INTRODUCTION

Collaborative Model Training

An approach that enables multiple organizations to train machine learning models on decentralized data, while maintaining data privacy and sovereignty.

Key Benefits

Enhanced privacy and security compared to traditional approaches, making it valuable in sectors with sensitive data (e.g. healthcare, building retrofitting financing).

BACKGROUND

Federated Learning

- Collaborative model training with decentralized data.
- Ensures privacy, governance, and controlled data access.
- Vertical Federated Learning
 - Utilizes distributed models for collaboration.
 - Handles heterogeneous features across clients.
 - Operates on a shared sample set among participants.





CONTRIBUTIONS

- Demonstrate Secure Collaborative Model Training with Vertical Federated Learning for the building energy consumption prediction use case.
- Use Microservices Architecture to allow for digital data exchange while ensuring privacy and contractual agreements.
- Use the FABRIC FAB Testbed to create a realistic, multisite networking slice and evaluate the implementation.

DYNAMOS – DYNAMICALLY ADAPTIVE MICROSERVICE-BASED OS

 Microservice-based: flexibility, scalability, and resilience by enabling independent deployment, management, and scaling of individual components

DYNAMOS Core

Dynamically creating and adapting the microservice chains according to policies and agreements

DYNAMOS Agents

The heterogeneous clients participating into the collaborative training



USE CASE

- Vertical Federated Learning for Building Energy Consumption Prediction
 - Harnessing data from diverse sources to predict energy usage.

Global Context:

- Buildings contribute 20-40% of energy consumption in developed countries, surpassing industrial and transportation sectors.
- HVAC systems are significant, accounting for:
 - 50% of building energy consumption.
 - 20% of total energy consumption in the USA.
- Source: "A Review on Buildings Energy Consumption Information")

USE CASE

Opportunities:

- Demand Forecasting: Predicts peak times to optimize usage and reduce waste.
- Retrofitting Advice: Identifies inefficiencies and suggests targeted energy-saving upgrades.
- Renewable Integration: Aligns energy consumption with sources like solar or wind.
- IoT and Smart Devices: Enables real-time adjustments for efficient energy use.

USE CASE

- Data providers:
 - EU Client 1: Building registry (area, floors, windows)
 - EU Client 2: Weather data (temperature, humidity)
 - US Client 1: Energy Provider (energy consumption)

	CONSUMPTION IS THE TARGET VARIABLE									TION IS THE Variable
ID	Area	Floors	Windows	ID	Temp	Humidity	Date	ID	Consumption	Date
1	130	1	10	1	20	80	2024-01-01	1	200	2024-01-08
4	50	1	3	4	24	90	2024-01-01	4	260	2024-01-08
5	75	2	4	5	16	90	2024-01-01	5	180	2024-01-08

MULTI-SITE FABRIC SLICE DEMO

- Demo focused on two international sites.
- Emulating a scenario with 3 inter-site clients:
 - 2 clients in EU AMST
 - 1 client in US DALL
- Model Aggregator is deployed at the US site.
- L2STS Networking, Transatlantic data transfers between sites.



MULTI-DOMAIN COLLABORATIVE MODEL TRAINING WITH VFL

DEMO: VFL FOR BUILDING ENERGY CONSUMPTION PREDICTION



CIENA TRAVELLING FABRIC NODE SC24 DEMO

- US site-only nodes hosted at the CIENA travelling FABRIC node, here at SC24!
- You can find us at the CIENA Booth 1940.
 - This way!





ANESTIS DALGKITSIS ALEXANDROS KOUFAKIS

UNIVERSITY OF AMSTERDAM

JORRIT STUTTERHEIM, ALEANDRO MIFSUD, PRIYANKA ATWANI, LEON GOMMANS, CEES DE LAAT, CHRYSA PAPAGIANNI, ANA OPRESCU

DISCUSSION FIND US: CIENA BOOTH 1940

