Work Package 4 – UNINA Federated Learning in Generative Models and Graph Neural Networks

Towards an Understanding of Artificial Intelligence via a transparent, open and explainable perspective - TUAIù
HORIZON-MSCA-2023-DN-01

Prof. Francesco Piccialli, University of Naples Federico II, Departments of Mathematics and Applications “R. Caccioppoli”, Italy

WP4 - Main Objectives

Objective 1: *Privacy-Preserving Federated Learning for Generative Models*

- **Focus:** Design innovative generative models (e.g., GANs, VAEs) that function within federated frameworks. Emphasize privacy by ensuring that sensitive data remains decentralized, maintaining high-quality data synthesis without compromising security.
- **Research Opportunity:** Explore adaptive privacy techniques, like differential privacy and homomorphic encryption, to safeguard data throughout the learning process. This approach could benefit sectors handling sensitive data, such as healthcare and finance.

Objective 2: *Advancement in Graph Neural Networks (GNNs) for Federated Learning*

- **Focus:** Improve GNN architectures to handle distributed, graph-structured data. This involves addressing data heterogeneity and ensuring that models maintain interpretability and efficiency in federated environments.
- **Research Opportunity:** Implement explainable AI (XAI) techniques specific to GNNs, enabling stakeholders to interpret complex model decisions, which could be crucial in critical sectors like smart infrastructure and IoT-based monitoring.

Objective 3: *Integrating NLP within Federated Learning and GNN Frameworks*

- **Focus:** Develop NLP models capable of decentralized language processing, where data privacy and security are critical. Utilize GNNs to enhance understanding of complex linguistic structures.
- **Research Opportunity:** Explore applications in smart healthcare and public safety, where federated NLP could analyze decentralized data (e.g., patient records, emergency response data) securely and transparently.

Deliverables

D4.1 - Federated AI Innovations Collection:

- **Content:** A comprehensive suite of generative models, GNNs, and NLP solutions designed for federated settings.
- **Potential Impact:** These models can serve as reference implementations for industry and academia, promoting privacy-aware AI adoption in domains like smart manufacturing and mobility.
- **Research Opportunity:** Publish and share these federated models on open platforms, fostering collaborative improvement and application of federated AI across various sectors.

D4.2 - Prototype Suite of Federated AI Systems and Evaluation Outcomes:

- **Content:** A set of federated AI prototypes rigorously tested in real-world scenarios, with performance metrics covering privacy, efficiency, and robustness.
- **Potential Impact:** Establishing a benchmark for federated AI performance, allowing stakeholders to assess model scalability and applicability across decentralized systems.
- **Research Opportunity:** Develop a robust evaluation framework focusing on privacy-preserving metrics, contributing valuable insights to the broader AI community regarding federated learning standards.

Tasks Overview

T4.1 - Federated Learning for Generative Models:

- **Activities:** Design prototype architectures that combine data synthesis with federated learning, ensuring scalability and data protection. Implement privacy-preserving mechanisms tailored for distributed data.
- **Research Opportunity:** Experiment with techniques like split learning to optimize the data flow between nodes, maintaining privacy without sacrificing model quality.

T4.2 - Advanced Studies on GNNs within Federated Learning:

- **Activities:** Adapt GNNs for decentralized data processing. Address challenges in synchronizing GNN updates across nodes, ensuring model efficiency.
- **Research Opportunity:** Investigate edge-computing approaches to minimize latency and improve the efficiency of GNNs in federated setups, relevant for real-time applications.

T4.3 - Integrating NLP with Federated Learning and GNNs:

- **Activities:** Develop NLP models that process language data across decentralized networks. Incorporate GNNs to capture contextual language relationships.
- **Research Opportunity:** Apply Federated NLP in sectors with sensitive data, such as legal tech or public sector, enabling secure, multi-node language data analysis.

T4.4 - Real-world Implementation and Testing of Frameworks:

- **Activities:** Deploy federated models in operational settings, evaluating real-world performance and refining models based on feedback.
- **Research Opportunity:** Partner with smart city or healthcare sectors to pilot and assess model effectiveness in managing decentralized data, contributing to policy and best-practice guidelines for federated AI deployment.

Doctoral Candidates and Projects (1/3)

Doctoral Candidate 1: *Privacy-Preserving Federated Generative Models for Decentralized Data Synthesis*

Objectives:

- Design and validate generative models within federated learning frameworks, emphasizing privacy preservation.
- Address challenges of decentralized data generation without compromising individual data privacy.

Expected Results:

- A robust federated generative model capable of synthesizing high-quality data while preserving data privacy.
- Development of protocols for deploying these models in privacy-sensitive domains, such as healthcare and finance.

Doctoral Candidates and Projects (2/3)

Doctoral Candidate 2: *Decentralized Graph Neural Networks: Adaptation, Training, and Interpretability in Federated Environments*

Objectives:

- Adapt and train Graph Neural Networks (GNNs) in federated learning settings to handle partitioned graph data.
- Enhance interpretability of GNNs within federated environments for better model transparency.

Expected Results:

- A prototype GNN suitable for federated environments, with minimal information leakage and high interpretability.
- Development of best practices and methodologies for applying GNNs in decentralized scenarios.

Doctoral Candidates and Projects (3/3)

Doctoral Candidate 3: *Enhancing NLP Capabilities through Federated Learning and GNNs*

Objectives:

- Integrate NLP within federated frameworks and leverage GNNs for processing complex language structures.
- Tackle challenges in decentralized NLP, focusing on privacy, data distribution, and model performance.

Expected Results:

- Federated NLP models capable of decentralized language processing while preserving data privacy.
- Comprehensive insights and best practices for implementing federated NLP in various applications.

Expected Outcomes

Federated Generative Models:

- **Outcome:** High-quality, privacy-preserving data generation models for decentralized settings.
- **Impact:** Enabling sensitive data applications (e.g., healthcare, finance) to leverage generative AI without compromising data privacy.
- **Research Opportunity:** Compare federated generative models against traditional centralized models to demonstrate equivalent performance with added privacy benefits.

Federated GNNs:

- **Outcome:** Efficient, interpretable GNNs adapted for federated environments.
- **Impact:** Support sectors needing transparent and decentralized data analysis, like transportation and energy networks.
- **Research Opportunity:** Publish methodologies for decentralized GNN training, contributing to the development of standards in federated GNN applications.

Federated NLP Models:

- **Outcome:** Decentralized NLP systems integrated with GNNs for complex data interpretation.
- **Impact:** Facilitate privacy-preserving language data processing for applications in legal, healthcare, and education.
- **Research Opportunity:** Develop and share guidelines on privacy-aware language processing, furthering ethical AI practices in NLP.

Innovation and Impact

Innovation in Privacy-Preserving AI:

- **Content:** WP4’s federated approaches combine AI’s data processing power with stringent privacy standards, meeting current demands for secure, decentralized systems.
- **Spinoff Potential:** These innovations have the potential to inspire new startups focusing on privacy-first AI solutions in various industries, from personalized healthcare to autonomous systems.

Impact on Sustainability and Efficiency in AI:

- **Content:** WP4 aims to create energy-efficient AI frameworks by leveraging federated architectures that minimize data transfers.
- **Research Opportunity:** Evaluate the energy consumption of federated vs. centralized models in GNNs and generative AI, contributing to the broader conversation on green AI.

Enabling Explainable AI (XAI) in Federated Contexts:

- **Content:** WP4 prioritizes model interpretability, especially for GNNs and NLP models in decentralized settings.
- **Potential Impact:** Facilitates compliance with AI transparency regulations (e.g., GDPR), opening opportunities for adoption in highly regulated sectors such as finance and healthcare.
- **Research Opportunity:** Innovate on explainability techniques specific to federated AI, offering insights that bridge technical implementation with ethical standards.

Collaboration with Project Partners

Research Partner Involvement: Outline cooperation with partners like ALMAWAVE SPA, CONFORM, UPM, SUT, NTNU.

Secondments & Training: Describe planned secondments for Doctoral Candidates to promote interdisciplinary skills and knowledge sharing across institutions.

Timeline and Milestones

Project Duration: 4 years, starting from October 1, 2024.

Recruitment of 3 Doctoral Candidates by June 2025