Multi-Agent Visualization for Explaining Federated Learning

Xiguang Wei[1], **Quan Li**[1], **Yang Liu**[1,2], **Han Yu**[4], **Tianjian Chen**[1] and **Qiang Yang**[1,3]
[1]AI Group, WeBank
[2]The Joint NTU-WeBank Research Centre of Eco-Intelligent Applications (THEIA)
[3]Department of Computer Science and Engineering, Hong Kong University of Science and Technology
[4]School of Computer Science and Engineering, Nanyang Technological University (NTU)
{xiguangwei,forrestli,yangliu,tobychen}@webank.com, han.yu@ntu.edu.sg, qyang@cse.ust.hk

Abstract

As an alternative decentralized training approach, Federated Learning enables distributed agents to collaboratively learn a machine learning model while keeping personal/private information on local devices. However, one significant issue of this framework is the lack of transparency, thus obscuring understanding of the working mechanism of Federated Learning systems. This paper proposes a multi-agent visualization system that illustrates what is Federated Learning and how it supports multi-agents coordination. To be specific, it allows users to participate in the Federated Learning empowered multi-agent coordination. The input and output of Federated Learning are visualized simultaneously, which provides an intuitive explanation of Federated Learning for users in order to help them gain deeper understanding of the technology.

1 Introduction

With advancement in data collection techniques and high-efficiency computing devices, data-driven machine learning has become the mainstream of engineering nowadays. Conventionally, in these data-driven systems, a centralized approach is adopted by traditional machine learning which requires the training data from different sources to be aggregated on a single machine or in a datacenter. This centralized training approach, however, is privacy-intrusive. In many applications, users have to sacrifice their privacy by sharing their personal data to train a better machine learning model. Recently, with several cases regarding privacy violation and harsher requirements by the General Data Protection Regulation (GDPR) by the European Union [Regulation, 2016; Voigt and Von dem Bussche, 2017], data privacy has become a hot issue in today's society. As an alternative decentralized training approach, Federated Learning (FL) enables users to collaboratively learn a machine learning model while keeping all the personal data that may contain private information on their local devices [Konečnỳ *et al.*, 2016; McMahan and Ramage, 2017; Yang *et al.*, 2019; Konečnỳ *et al.*, 2015; Bonawitz *et al.*, 2019]. In such a case, users can benefit from a well-trained machine learning model without sharing their sensitive personal data.

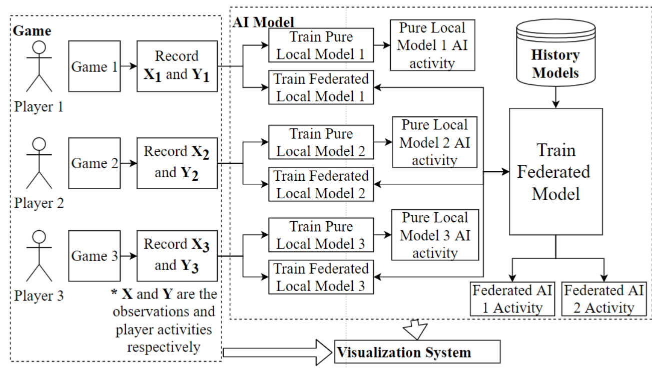

Figure 1: System pipeline. Users are involved in the car racing game play. Their view observations and actions are used to train two kinds of AI models, which are local AI and FL. Both AI models can drive one or more cars to race in the game. Their data are then fed into the visualization system for further demonstration.

Despite its wide applicability in industrial, medical, and financial scenarios [Liu *et al.*, 2018; Huang and Liu, 2019; Huang *et al.*, 2018; Hankz Hankui Zhuo and Lin, 2019; Kumar *et al.*, 2017], FL has its own problems. The most significant one is the lack of transparency behind their behaviors, which leaves users, e.g., collaboration partners and customers, with little technical background in this area very confused. The concerns about the non-transparent nature of FL have hampered its wider adoption [Du *et al.*, 2018].

In this work, we showcase a platform for intuitive demonstration and explanation of how a typical FL system works. To be specific, we built a multi-agent visualization platform to illustrate what is FL and how it supports the privacy-preserving multi-agent coordination. The platform consists of three parts: 1) a controllable racing game based on multi-agent box cars; 2) AI models running behind the game, and 3) an intuitive visualization system for the explanation of the training and inference processes of this game.

2 Introduction to FL

FL was first proposed in by Google as "*a specific category of distributed machine learning approaches which trains machine learning models using decentralized data residing on end devices such as mobile phones*" [McMahan and Ramage, 2017]. Lenovo built hardware to simulate industrial pro-

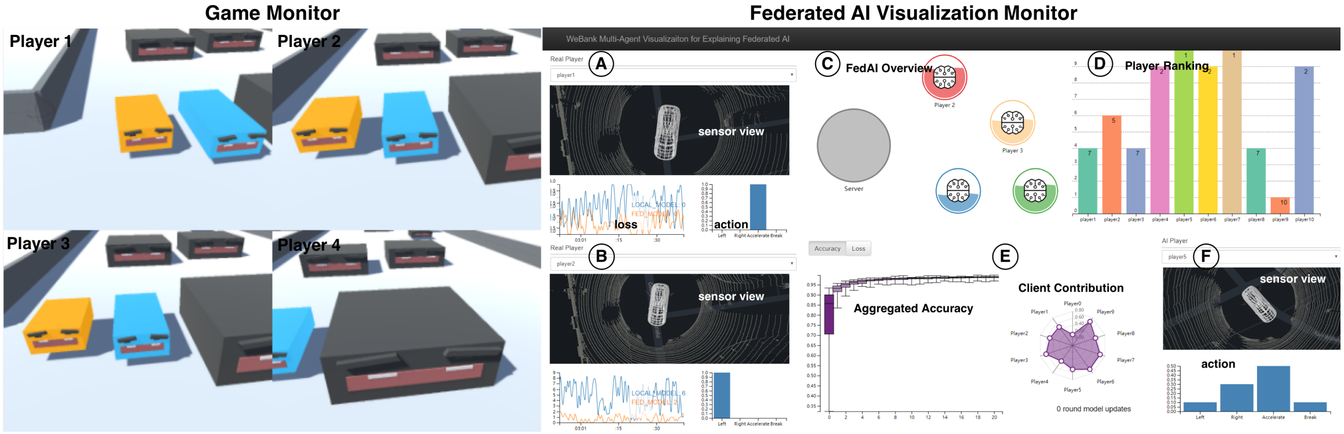

Figure 2: Our multi-agent visualization system to explain FL. The left figure shows the real-time gameplay of four cars controlled by real players, and the right figure demonstrates our visualization system, which consist of: (A,B) Player View which includes a sensor view based on point cloud data, a loss function view and an action view; (C) a FL Overview; (D) Player Ranking View; (E) Aggregated Accuracy and Client Contribution View, and (F) AI View.

cesses in factories [Rojek, 2018], and the objective was to predict how the pressure inside the hardware. Cloudera Fast Forward Labs showcased an interactive simulation prototype, Turbofan Tycoon that leverages visualizations to communicate the advantages of Federated model which makes more accurate predictions about when a turbofan will fail [Mike, 2018]. However, one challenge they faced is that the case of preventative maintenance is "*not necessarily everyone's idea of an exciting or approachable topic*". They further indicated that if FL is simulated in a video game scene, lots of niche technical details can be a source of entertainment, which also motivates our work.

3 The Demonstration System

3.1 Racing Game

We designed simple rules and interactive inputs to facilitate user involvement in the racing game. To be specific, in the box car racing game, each player can control a box-like car and the one who first reaches the finish line wins. During the racing, randomly generated obstacles will appear in the racing track to slow down the racing. In each round of the game, four cars are controlled by real players, while the other six are controlled by AI models. During each game play, as shown in Fig.2, we use X and Y to denote the observations and actions of each player, respectively, which will be recorded for further training of AI models. These data will be fed into the subsequent visualization system so that players can have an intuitive overview and understanding of the generated input and output to feed the AI models.

3.2 AI Models

We train AI models to control cars in the game play. A multilayer perceptron (MLP) is leveraged to generate the AI models. As shown in Fig. 1, given the recorded data for each player, two variants of AI models, i.e., a local AI model and a federated local model, are generated and trained for each player. The local model leverages purely the recorded data of the corresponding player, while the federated local model keeps updating itself by communicating model information with all clients following the FL protocols.

3.3 Visualization

We develop three main visualizations to demonstrate the local view and the performance of local models in terms of a local model and a federated local model for each player, the federated aggregation model, and the AI model activities, respectively. Specially, we design 1) a player view which consists of a sensor view (point cloud data based to visualize the observation and surroundings of the car), a loss view to illustrate the loss function and an action view to show the current actions of the player; 2) a FL overview to illustrate model updates and communications between clients and the server, a player ranking view and an aggregated accuracy view to show the performance of FL and contributions of each client, and 3) an AI view which is similar to the player view but it shows the observation and the actions of AI-controlled cars. Users can select any two real players in the player view and their data will be simultaneously visualized in the system. Similarly, we can select any one of the other six AI model-controlled cars in the AI view for observation.

4 Conclusions and Future Work

The system demonstrated in this paper is a promising educational tool for FL. In the future, we plan to generalize this framework to other FL scenarios and incorporate hardware components for better illustration.

Acknowledgements

This research is supported, in part, by the Nanyang Assistant Professorship (NAP) and the Joint NTU-WeBank Research Centre of Eco-Intelligent Applications (THEIA), Nanyang Technological University, Singapore.

References

[Bonawitz *et al.*, 2019] Keith Bonawitz, Hubert Eichner, Wolfgang Grieskamp, Dzmitry Huba, Alex Ingerman, Vladimir Ivanov, Chloe Kiddon, Jakub Konecny, Stefano Mazzocchi, H Brendan McMahan, et al. Towards federated learning at scale: System design. *arXiv preprint arXiv:1902.01046*, 2019.

[Du *et al.*, 2018] Mengnan Du, Ninghao Liu, and Xia Hu. Techniques for interpretable machine learning. *arXiv preprint arXiv:1808.00033*, 2018.

[Hankz Hankui Zhuo and Lin, 2019] Qian Xu Qiang Yang Hankz Hankui Zhuo, Wenfeng Feng and Yufeng Lin. Federated reinforcement learning. *arXiv preprint arXiv:1901.08277*, 2019.

[Huang and Liu, 2019] Li Huang and Dianbo Liu. Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *arXiv preprint arXiv:1903.09296*, 2019.

[Huang *et al.*, 2018] Li Huang, Yifeng Yin, Zeng Fu, Shifa Zhang, Hao Deng, and Dianbo Liu. Loadaboost: Loss-based adaboost federated machine learning on medical data. *arXiv preprint arXiv:1811.12629*, 2018.

[Konečnỳ *et al.*, 2015] Jakub Konečnỳ, Brendan McMahan, and Daniel Ramage. Federated optimization: Distributed optimization beyond the datacenter. *arXiv preprint arXiv:1511.03575*, 2015.

[Konečnỳ *et al.*, 2016] Jakub Konečnỳ, H Brendan McMahan, Felix X Yu, Peter Richtárik, Ananda Theertha Suresh, and Dave Bacon. Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*, 2016.

[Kumar *et al.*, 2017] Saurabh Kumar, Pararth Shah, Dilek Hakkani-Tur, and Larry Heck. Federated control with hierarchical multi-agent deep reinforcement learning. *arXiv preprint arXiv:1712.08266*, 2017.

[Liu *et al.*, 2018] Yang Liu, Tianjian Chen, and Qiang Yang. Secure federated transfer learning. *arXiv preprint arXiv:1812.03337*, 2018.

[McMahan and Ramage, 2017] Brendan McMahan and Daniel Ramage. Federated learning: Collaborative machine learning without centralized training data. *Google Research Blog*, 3, 2017.

[Mike, 2018] Mike. Federated learning: distributed machine learning with data locality and privacy. https://blog.fastforwardlabs.com/2018/11/14/federated-learning.html, 2018.

[Regulation, 2016] Protection Regulation. Regulation (eu) 2016/679 of the european parliament and of the council. *REGULATION (EU)*, 679, 2016.

[Rojek, 2018] Marcin Rojek. Devices learning from each other? see it live this september at ai summit in san francisco! https://medium.com/@marcrojek/devices-learning-from-each-other-see-it-live-this\-september-at-ai-summit-in-san-francisco-9ac6462d7f8e, 2018.

[Voigt and Von dem Bussche, 2017] Paul Voigt and Axel Von dem Bussche. The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 2017.

[Yang *et al.*, 2019] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2):12, 2019.