사용자 편의성 제고를 위한 SplitNN 프레임워크 컨테이너화

심채린[1], 조연정[1], 이명준[1], 윤태현[2]
[1]체인트리 주식회사, [2]한국전자통신연구원
[1]{chaelin0520, jeong7323}@gmail.com, [1]mjlee@cicweb.ulsan.ac.kr, [2]thyoon0820@etri.re.kr

Containerization of the SplitNN Framework for Enhancing User Convenience

Sim Chae Lin[1], Jo Yeon Jeong[1], Lee Myung Joon[1], Tae Hyun Yoon[2]
[1]Chaintree Inc., [2]Electronics and Telecommunications Research Institute

요 약

본 논문은 SplitNN 기반 협업 학습의 신뢰성 및 무결성을 지원하는 블록체인 기반 실시간 모니터링 시스템인 Train4U의 복잡한 설정 문제를 해결하고자, Docker 기반의 자동화 실행 환경을 설계하였다. 제안된 구조는 환경 설정, 블록체인 연동, 학습 실행 등을 자동화하여, IT 인프라에 익숙하지 않은 제조기업도 쉽게 활용할 수 있도록 지원한다. 이를 통해 Train4U의 산업 현장 적용성을 높이고, AI 협업 학습 시스템의 실효성을 강화할 수 있다.

Ⅰ. 서론

최근 제조, 금융, 의료 등 다양한 산업군에서 인공지능(AI) 기술을 도입하려는 시도가 활발히 이루어지고 있다. 특히 민감한 데이터를 다루는 환경에서는 데이터의 기밀성을 보장하면서도 효과적인 딥러닝 학습을 수행할 수 있는 기술이 요구된다[1]. 이러한 배경 속에서, 데이터를 외부로 노출하지 않고도 분산 환경에서 협업 학습이 가능한 스플릿 뉴럴 네트워크(Split Neural Network, SplitNN) 기법이 주목받고 있다. SplitNN[2]은 딥러닝 모델을 상부와 하부로 분리하여, 데이터 소유자와 AI 서비스 제공자가 각자의 영역에서 모델을 학습함으로써 원본 데이터를 보호하면서도 공동 학습이 가능하다. 이러한 기술적 요구를 반영하여, 기존 연구에서는 블록체인 기반의 실시간 모니터링 시스템인 Train4U[3]를 제안하였다. Train4U는 학습 과정을 블록체인에 기록하고, 중간 결과에 대한 위·변조 방지와 무결성 검증 기능을 통해 학습의 신뢰성과 투명성을 확보할 수 있도록 설계되었다.

Train4U는 제조기업과 AI 서비스사가 공동으로 활용할 수 있는 협업 학습 프레임워크로, 보안성과 신뢰성을 동시에 충족하는 견고한 구조를 갖추고 있다. 그러나 복잡한 시스템 설정 절차는 IT 전문기술이 부족한 제조기업에게 높은 진입 장벽으로 작용하며 이로 인해 Train4U 기반 협업 학습 프레임워크의 확산과 활성화가 저해되고 있다.

이에 본 연구에서는 사용자 피드백을 바탕으로, 복잡한 설정 없이도 Train4U를 손쉽게 실행할 수 있도록 컨테이너 기반 자동화 실행 환경을 설계하였다. 각 구성 요소는 Docker[4] 이미지로 통합되었으며, 환경 변수 자동 설정, 네트워크 구성, 블록체인 및 데이터베이스 연동 절차를 자동화하여, 양측 사용자 모두가 최소한의 기술 개입만으로 시스템을 운용할 수 있도록 구성하였다. 이를 통해 Train4U의 초기 설정 부담을 경감시키고, 제조기업과 AI 서비스사 간 협업 학습의 현장 도입 가능성을 높이고자 한다.

Ⅱ. 본론

본 연구에서는 제조기업과 AI 서비스사가 협업하여 SplitNN 기반 트레이닝을 보다 쉽게 수행할 수 있도록, Train4U 시스템의 제조기업 측 실행 환경을 컨테이너 기반으로 구성하였다. 기존 Train4U는 개발사가 초기 환경 설정을 지원하지만, 실제 운용 중 설정을 변경하거나 문제를 해결하는 과정에서 IT 인력이 부족한 제조기업 입장에서는 지속적인 활용에 어려움을 겪는 사례가 존재하였다. 특히 블록체인 연동 설정, 데이터 흐름 구성, 스크립트 수정 등의 작업은 비전문 사용자에게 높은 기술적 부담으로 작용한다.

이를 개선하기 위해 본 연구에서는 제조기업 측 환경을 Docker 기반 이미지로 구현하고, 자동화된 실행 스크립트를 함께 제공하여 시스템 실행 과정을 단순화하였다. 컨테이너 이미지에는 SplitNN 하위 모델 실행 코드, 블록체인 클라이언트 구성, 분산 데이터베이스 연결 정보 등이 포함되어 있으며, 별도의 복잡한 설정 없이도 실행이 가능하도록 설계되었다. 전체 구동은 docker-compose.yml[5]을 기반으로 하며, 초기 환경 변수 설정부터 블록체인 네트워크 구성, 학습 시작까지의 과정을 자동화된 흐름으로 처리한다. 이미지에 포함된 주요 구성 요소는 그림 1에 정리하였다.

구성 요소	설명
OS 환경	Ubuntu 20.04 기반 컨테이너 운영체제
Python	Python 3.9 버전 기반 딥러닝 실행 환경
PyTorch	PyTorch 1.13.1, CUDA 11.7 지원 포함
블록체인 SDK	Hyperledger Fabric Node SDK 내장 (블록체인 연동 처리)
DB 클라이언트	Cassandra Python Driver 포함 (분산 DB 연결용)
웹 프레임워크	Django 4.2 + Gunicorn + Nginx 기반 웹 인터페이스 제공
자동화 스크립트	Shell 기반 자동 실행 스크립트 (‘start_all.sh‘, ‘init.sh‘)
환경 설정	‘.env‘ 및 설정 파일 기반 자동 변수 로딩
모니터링 UI	Django 기반 설정, 실행, 모니터링 기능 포함 웹 UI

그림 1. Train4U Docker 이미지 구성 요소 및 설명

자동화된 실행은 단일 스크립트 실행만으로 초기화 및 학습 준비 환경이 자동으로 구성되도록 설계되었다. 사용자가 start_all.sh와 같은 통합 스크립트를 실행하면, 사전에 구축한 Docker 이미지 기반으로 컨테이너가 구동되고, 그 내부에서 환경 변수 설정, 블록체인 네트워크 및 분산 데이터베이스 연결 등의 절차가 순차적으로 처리된다. 모든 설정과 실행은 컨테이너 내부에서 자동화되며, 사용자 개입은 최소화된다. 자동화된 실행 프로세스의 전체 흐름은 그림 2의 시퀀스 다이어그램을 통해 확인할 수 있다.

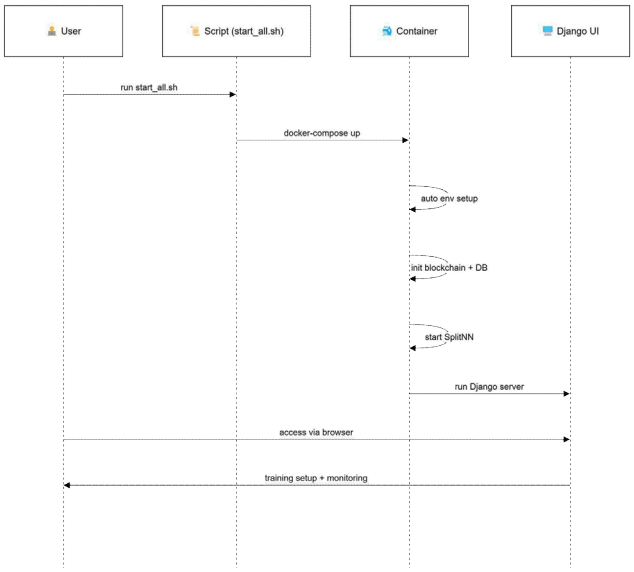

그림 2. Train4U 자동화 실행 흐름 시퀀스 다이어그램

초기 셋업이 완료되면, Django[6] 기반의 사용자 인터페이스(UI)가 실행되어 제조기업 사용자는 로컬에 존재하는 학습 데이터를 기반으로 설정을 조정하고, SplitNN 학습 실행 및 모니터링을 웹 환경에서 손쉽게 수행할 수 있다. 이후 시스템은 자동으로 AI 서비스사 측과 연결되어 분산 협업 학습을 시작한다. 이를 통해 기술적 전문성이 없는 제조기업에서도 안정적으로 시스템을 운용하는 것이 가능하다.

이러한 컨테이너 기반 자동화 구조는 기존의 수동 구성 방식에 비해 시스템 실행에 필요한 IT 인력 확보 부담을 줄이고, 설정 절차의 단순화 및 실행의 표준화를 통해 운영 안정성을 높이며, 동일 이미지 재사용을 통해 테스트 및 확장 환경에서 높은 재현성을 확보하는 등의 이점을 제공한다. 따라서 제안된 방법은 제조기업이 기술 인프라에 대한 부담 없이 AI 학습에 참여할 수 있도록 지원하며, AI 서비스 기업과의 협업을 안정적이고 효율적으로 운영할 수 있는 실질적인 기반이 될 수 있다.

Ⅲ. 결론

본 연구에서는 SplitNN 기반 협업 학습 시스템인 Train4U의 초기 설정 및 실행 절차에서 발생하는 사용자 불편을 해소하기 위해, Docker 기반의 자동화된 실행 환경을 제안하였다. 제조기업 측 실행 환경을 단일 이미지 기반으로 구성하고, 실행 스크립트와 UI를 포함한 자동화 구조를 적용함으로써 복잡한 설정 없이도 손쉽게 학습을 수행할 수 있도록 설계하였다.

제안된 구조는 특히 IT 인프라가 부족한 제조기업에서도 안정적인 실행이 가능하도록 구현되었으며, 설정의 단순화, 실행의 재현성, 초기 진입 장벽 완화 측면에서 높은 실용성을 갖는다. 향후에는 실행 프로세스의 자동화 범위를 확대하고, 사용자 환경에 따른 맞춤형 설정 및 시스템 안정성 향상을 중심으로 기능을 고도화함으로써, Train4U의 실용성과 적용 범위를 지속적으로 확장하고자 한다.

ACKNOWLEDGMENT

본 논문은 울산시-ETRI 2차 공동협력사업의 일환으로 수행되었음. [25AB1600, 제조 혁신을 위한 주력산업 지능화 기술 개발 및 산업현장에서의 사람-이동체-공간 자율협업지능 기술 개발]

참 고 문 헌

[1] P. Vepakomma, O. Gupta, T. Swedish, and R. Raskar, “Split learning for health: Distributed deep learning without sharing raw patient data,” arXiv preprint arXiv:1812.00564, 2018.

[2] MIT Media Lab, "MIT Media Lab's Split Learning: Distributed and collaborative learning," Available: http://splitlearning.mit.edu, accessed on November 10, 2021.

[3] J. Y. Jeong, C. L. Sim, Y. G. Yoon, T. H. Yoon, and M. J. Lee, “A study on the trustworthy split learning execution methodology based on blockchain systems,” Proceedings of the Korean Institute of Communications and Information Sciences Conference, Gyeongbuk, Korea, Nov. 2023.

[4] Docker, https://docs.docker.com/

[5] C. Boettiger, “An Introduction to Docker for Reproducible Research,” ACM SIGOPS Operating Systems Review, vol. 49, no. 1, pp. 71–79, Jan. 2015.

[6] Django, https://docs.djangoproject.com/