연합 학습을 통한 사기 탐지 시스템

19th Jiyun Kang, 20th Chaewon Kim, Yujin Lee, 21th Suhwan Kim
KUBIG(Korea University Data Science & AI Society), Korea University
(k-novin01, kcherry, snoapple, ksh61148)@korea.ac.kr

고려대학교 KOREA UNIVERSITY

KUBIG DATA SCIENCE & AI 8th CONFERENCE

Motivation

- 금융도메인의 이상거래탐지(Fraud Detection) 모델 개발에는 대규모 금융데이터가 필요
- Challenge 1 : 개별 금융기관의 data만으로 학습된 모델은 해당 기관의 data 특성에 overfitting되어 일반화 성능 확보가 어려움 ▶ 기관 간의 연합학습이 필요
- Challenge 2 : 금융 데이터는 Sensitive Data로 각 금융기관 간의 데이터 공유가 현실적으로 어려움
- **Solution: 연합 학습(Federated Learning)**

▶ 데이터를 직접 공유하지 않고 각 기관에서 독립적으로 학습한 모델의 결과 Parameter만을 취합하여 공유하는 분산 학습 방식

Key Idea

- Phase 1: 이상 거래 탐지 모델 고도화

Centralized setting(즉, 중앙 단일 기관(Single Client) 상황 가정) 이상 거래 탐지 모델 개발 → FL 성능 비교의 Upper bound 기준점

- Phase 2: 연합학습 프레임워크 적용

고도화된 Fraud Detection 모델을 각 금융기관에 분산 배포하고, 데이터 프라이버시를 보장하면서 협력적으로 학습하여 단일 글로벌 모델(Global Model)을 구축

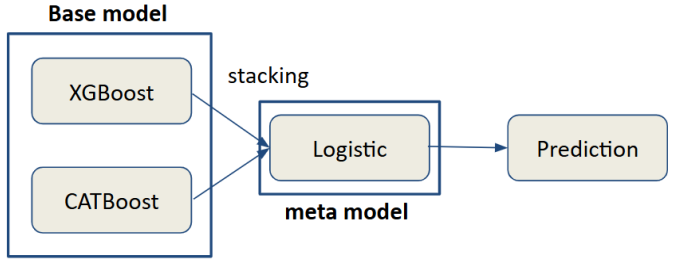
▶ **핵심 목표: 성능 격차 최소화**

연합 학습을 통해 생성된 최종 글로벌 모델의 성능이, 모든 데이터가 하나의 서버에 모여 있다고 가정한 이상적인 중앙 집중식 학습(Centralized Setting) 모델의 성능에서 얼마나 적게 하락하였는지, 최대한 근접하도록

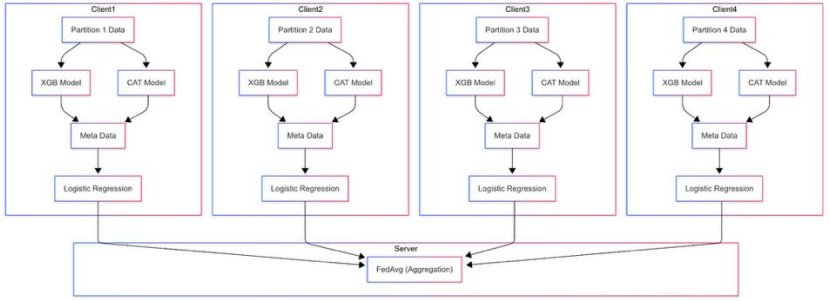
Method

- **Dataset**
  - **FSI AIxData Challenge 2024에서 공개한 dataset**
  - **Account(송금인 정보) + Transaction (거래 정보)**
- **Phase 1 : 이상거래탐지 모델 고도화**

▶ **Centralized learning에서의 최종 성능 : 0.83 (binary f1 기준)**



- **Phase 2 : 연합학습 프레임워크 적용**

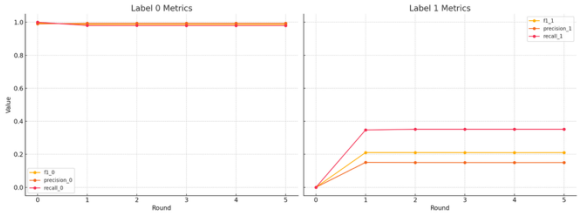


기존 **중앙 집중형 Stacking 구조**를 연합학습 **분산형 구조로 구현**

- 각 클라이언트는 **Base 모델 학습 및 메타데이터 생성**
- 서버는 **Logistic Regression 기반 메타 러너**를 연합 방식으로 **통합 학습**

Result

- **Centralized Learning**
  Precision 0. 96, Recall 0.73, F1 score 0.83
- **Federated Learning (round 5)**
  Precision 0.15, Recall 0.35, F1 score 0.21



Conclusion

- **FL 이 CL보다 비열등성 입증 실패**

원인 1) 클래스 불균형
원인 2) Base model 간 불일치
- 클라이언트 간의 서로 다른 모델 사용으로 인해 meta learner의 일관된 학습이 어려움

원인 3) 스태킹 모델의 적용

- **Future Work**

1. **Base model 통일**
2. **Server의 base model 보완**
3. **Aggregation 방식의 개선**