

불균형 금융 데이터 환경에서의 머신러닝 기반의 이상 거래 탐지 성능 분석

이채호, 홍용근, 정관수*, 노학균

대전대학교, *교원대학교

{chaeho0902@gmail.com, yghong@dju.kr, hg.roh@dju.kr}, *ksjung@knue.ac.kr

Performance Analysis of ML-based Bank Fraud Detection on Imbalanced Data

Yi Chae Ho, Yong-Geun Hong, Kwansoo Jung*, Hakgyun Roh

Daejeon University, *Korea National Univ. of Education

요약

디지털 금융 환경의 확산과 비대면 금융 서비스의 활성화로 인해 금융 사기 발생 빈도와 피해 규모가 지속적으로 증가하고 있다. 기존의 규칙 기반 사기 탐지 시스템은 정적인 규칙에 의존하여 새로운 사기 유형에 대한 대응이 어렵고, 과도한 오탑(False Positive)과 미탐(False Negative) 문제를 야기한다. 본 연구에서는 실제 은행 계좌 개설 데이터를 기반으로 생성된 대규모 불균형 데이터셋을 활용하여 금융 사기 탐지에 적합한 머신러닝 모델을 비교·분석하고, 운영 환경에 최적화된 임계값 전략을 제안한다. Random Forest, XGBoost, Light GBM 모델을 중심으로 성능과 학습 효율을 비교한 결과, 부스팅 계열 모델이 높은 재현율과 우수한 학습 속도를 보이며 실제 금융 사기 탐지 환경에 가장 적합함을 확인하였다.

I. 서론

디지털 금융 기술의 발전은 금융 서비스의 접근성과 편의성을 크게 향상시켰으나, 특히 비대면 계좌 개설 단계에서의 사기 행위는 대포통장 개설이나 자금세탁 등 후속 범죄로 이어질 가능성이 높아 금융기관의 핵심 리스크 요인으로 지목된다. 현재 대다수 금융기관은 규칙 기반(Rule-based) 시스템을 운용하고 있으나, 이는 사전 정의된 정적 규칙에 의존하기 때문에 지능화된 신규 및 변종 사기 유형에 실시간으로 대응하는 데 구조적인 한계를 가진다. 또한, 경직된 판단 기준으로 인해 정상 거래를 차단하는 오탑(False Positive)과 실제 사기를 놓치는 미탐(False Negative) 문제가 빈번하여, 데이터 기반의 고도화된 탐지 기법 도입이 요구되고 있다. 인공지능 기반의 다양한 이상 거래 탐지가 대안으로 주목받고 있으나, 실제 금융 데이터 환경에 이를 적용하기 위해서는 극심한 클래스 불균형(Class Imbalance) 문제를 해결해야 한다. 정상 거래가 전체 데이터의 대부분을 차지하는 환경에서 일반적인 모델을 학습 시킬 경우, 알고리즘이 다수 클래스인 정상 패턴 위주로 최적화되어 소수 클래스인 사기 거래에 대한 재현율(Recall)이 현저히 저하되는 현상이 발생한다. 사기 거래의 미탐은 금융기관에 막대한 금전적·사회적 손실을 초래하므로, 단순한 정확도 향상을 넘어 불균형 데이터를 효과적으로 학습하고 사기 탐지율을 보장할 수 있는 정교한 모델링이 요구된다.

본 연구는 대규모 불균형 데이터 환경에서도 견고한 성능을 발휘하는 머신러닝 기반 이상 거래 탐지 프레임워크를 제안한다. 이를 위해 실제 은행 계좌 개설 로그 기반의 벤치마크 데이터셋을 활용하여, 데이터 수준의 불균형 해소를 위해 전처리 과정을 통해서, 사기 데이터의 학습 비중을 보완한다. 또한 비교 분석을 위해 대표적인 양상블 모델을 선정하여 성능을 비교 및 검증한다. 특히 본 연구는 단순 예측 정확도 평가를 지양하고, 사기 적발의 핵심 지표인 재현율을 극대화하기 위한 임계값 최적화 전략을 수행하여 적용가능성을 확인하고자 한다.

1. 데이터셋 및 전처리

실제 금융 환경의 특성을 반영하기 위해서 BAF(Bank Account Fraud) 데이터셋을 활용하였다[1]. 은행 계좌 개설 로그를 기반으로 CTGAN(Conditional Tabular GAN) 기법을 적용하여 생성되었으며, 전체 데이터는 총 1,000,000건의 인스턴스로 구성되어 있으나, 이 중 사기(Fraud) 계좌의 비율은 약 1.1%에 불과하여 극심한 클래스 불균형(Class Imbalance) 특성을 보인다.

데이터 전처리 과정에서는 다음과 같이 3단계 프로세스로 수행하였다.
(1) 특성 선별 및 가공 단계: 소득 수준(income), 이름-이메일 유사도(name_email_similarity), 단기·중기 거래 속도(velocity_6h, 24h) 등 금융 사기 탐지에 높은 설명력을 가지는 인구통계학적 정보와 행동 기반 패턴을 주요 피처로 선정하였다.

(2) 데이터 변환 단계: 범주형 변수에 One-Hot Encoding을 적용하여 수치화하고, 수치형 변수에는 정규화(Normalization)를 수행하여 변수 간 스케일 차이에 따른 학습 불안정성을 해소하였다.

(3) 데이터 분할 및 불균형 해소 단계: 전체 데이터를 학습용(80%)과 테스트용(20%)으로 분리하였고, 계층적 샘플링(Stratified Sampling)을 통해 원본의 클래스 분포를 유지하였다. 특히, 모델이 다수 클래스인 정상 거래에 편향되는 것을 방지하기 위해 학습 데이터에 한하여 SMOTE (Synthetic Minority Over-sampling Technique) 기법을 적용하였다[2]. SMOTE는 소수 클래스인 사기 데이터를 인위적으로 합성 및 증식시켜 데이터 불균형 문제를 구조적으로 보완함으로써, 모델이 정상 거래에 편향되지 않고 사기 패턴을 효과적으로 학습할 수 있도록 돋는 데이터 증강 기법이다. 본 연구에서는 소수 클래스인 사기 데이터를 인위적으로 합성·증식시켜 데이터 불균형 문제를 구조적으로 보완하고 사기 패턴에 대한 학습 기회를 강화하였다.

2. 모델 설계 및 하이퍼파라미터 설정

불균형 데이터 환경에서의 사기 탐지 성능을 비교·분석하기 위해 트리 기반 양상블 머신러닝 모델을 비교 모델로 선정하였다. 비교 모델로는 배깅(Bagging) 방식의 Random Forest와 부스팅(Boosting) 방식의

II. 본론

XGBoost, LightGBM을 채택하였다[3]. 이들 모델은 금융 사기 탐지 분야에서 높은 예측 성능과 안정성을 보이는 대표적인 알고리즘으로, 실제 금융 시스템에서의 적용 가능성이 높다는 공통점을 가진다.

Random Forest는 다수의 결정 트리를 독립적으로 학습시키는 배깅 기반 양상을 기법으로, 데이터 분산을 줄이고 과적합에 강건한 특성을 가진다. 그러나 트리들이 순차적 보완 없이 독립적으로 학습되기 때문에, 극심한 클래스 불균형 환경에서는 소수 클래스에 대한 민감도가 상대적으로 낮을 수 있으며, 대규모 데이터셋에서는 학습 시간과 계산 자원이 많이 소요된다는 한계를 가진다.

XGBoost는 Gradient Boosting 기반 모델로, 이전 학습기의 오차를 순차적으로 보완하는 방식으로 모델을 학습한다. L1·L2 정규화, 결측치 자동 처리, 학습 속도 최적화 기법 등을 제공하여 높은 일반화 성능을 보이며, 특히 불균형 데이터 환경에서 가중치 조정을 통해 사기 거래와 같은 소수 클래스 탐지 성능을 효과적으로 향상시킬 수 있어 금융 사기 탐지 분야에서 활용되고 있다.

LightGBM은 Histogram 기반 분할 방식과 Leaf-wise 트리 성장 전략을 적용한 Gradient Boosting 모델로, 대규모 데이터 환경에서도 빠른 학습 속도와 우수한 메모리 효율성을 제공한다. 특히 테이블 형태의 금융 데이터에서 높은 예측 성능을 보이며, 실시간 또는 준실시간 사기 탐지와 같은 운영 환경에 적합한 특성을 가진다.

모든 모델은 공정한 성능 비교를 위해 트리 개수(n_estimators)를 500으로 통일하여 학습하였다. 또한 부스팅 계열 모델(XGBoost, LightGBM)의 경우, 클래스 불균형 문제를 완화하고 사기 탐지 민감도를 강화하기 위해 소수 클래스에 대한 가중치 조정(weighting) 전략을 적용하였다.

Table 1. 하이퍼파라미터 설정

유형	RandomForest	XGBoost	LightGBM
결정트리수	500	500	500
학습률	-	0.05	0.05
트리 깊이	Unlimited	7	-1(Leaf-wise)
클래스 가중치	Balanced	scale_pos_weight	scale_pos_weight
최적화 기준	Gini Impurity	Logloss	Logloss

III. 실험 결과

1. 평가 지표 및 임계값 설정

금융 사기 탐지 문제는 실제 사기 거래를 정상으로 오판(False Negative)할 경우 발생하는 금전적·사회적 손실이 매우 크다는 특성을 가진다. 따라서 본 연구에서는 일반적인 분류 문제에서 주로 활용되는 정확도(Accuracy)나 정밀도(Precision)보다, 사기 거래를 얼마나 효과적으로 탐지하는지를 나타내는 재현율(Recall)을 핵심 평가지표로 선정하였다. 이를 위해 모델이 출력한 사기 확률(Probability)을 이진 분류 결과로 변환하는 과정에서 기본 임계값 0.5가 아닌 0.2으로 하향 조정하였다. 해당 임계값 전략은 모델의 탐지 민감도를 높여 사기 거래 누락을 최소화하기 위한 것으로, 금융 사기 탐지의 비대칭적 오류 비용 구조를 반영한 설정이다.

2. 모델별 성능 비교 분석

Random Forest, XGBoost, LightGBM의 성능을 동일한 데이터 및 임계값(0.2) 조건에서 비교한 결과는 다음과 같다. 임계값 조정 후 Random Forest는 재현율 37.9%로 나타났으며, SMOTE 적용에도 불구하고 극심한 클래스 불균형 환경에서 사기 거래 탐지에 한계를 보였다. 또한 오탐(False Positive)이 상대적으로 많이 발생하여 실무 적용 측면에서 제약이 확인되었다. 반면 부스팅 기반 모델인 XGBoost는 재

현율 88.2%를 기록하여 Random Forest 대비 사기 탐지 성능이 크게 향상되었다. LightGBM은 재현율 90.1%로 세 모델 중 가장 높은 탐지 성능을 달성했으며, XGBoost와 유사한 수준의 재현율을 유지하면서도 오탐을 비교적 효과적으로 억제해 성능과 안정성의 균형이 우수했다. 이러한 결과는 LightGBM의 Histogram 기반 분할 및 leaf-wise 성장 방식이 대규모 불균형 데이터에서 효율적인 학습에 기여한 것으로 볼 수 있다.

Table 2. 비교 모델 별 재현율/학습시간

	RandomForest	XGBoost	LightGBM
임계값(Threshold)	0.2	0.2	0.2
재현율(Recall, %)	37.9	88.2	90.1
학습시간(second)	1,365	225	48

실제 금융 운영 환경은 대규모 트랜잭션이 실시간으로 발생하기 때문에, 모델의 탐지 성능 못지 않게 학습 및 재학습(Re-training) 속도가 핵심적인 운영 지표로 작용한다. 본 연구에서 각 모델별 학습 시간을 비교 분석한 결과, Random Forest는 약 22분 15초가 소요된 반면, XGBoost는 약 3분 45초, LightGBM은 48초 미만의 학습 시간을 기록하며 모델 간 현저한 효율성 차이를 나타냈다. 특히 LightGBM은 Random Forest 대비 약 27.8배의 빠른 학습 효율성을 보여, 급변하는 사기 패턴에 즉각적으로 대응하기 위한 빈번한 재학습과 신속한 배포가 요구되는 금융 운영 환경에서 중요한 강점으로 작용한다. 종합적으로 LightGBM은 비교 모델 중 가장 높은 사기 탐지율(재현율)과 동시에 운영 효율성(학습시간)을 동시에 확보함으로써, 대규모·극심한 불균형 금융 데이터 환경에서 효과적인 모델임을 확인하였다.

IV. 결론

본 연구는 비대면 금융 서비스 확대로 증가하는 금융 사기를 효과적으로 차단하기 위해, 머신러닝 기반 이상거래 탐지 모델의 성능을 비교·분석하는 것을 목적으로 한다. 극심한 클래스 불균형 문제를 완화하기 위해 SMOTE를 적용하여 학습 데이터를 증강하고, 사기 탐지의 핵심 지표인 재현율(Recall)을 극대화하기 위해 분류 임계값 최적화 전략을 도입하였다. 실험 결과, Random Forest 대비 XGBoost와 LightGBM 등 부스팅 계열 모델이 불균형 데이터 환경에서 더 높은 탐지 성능과 빠른 학습 효율을 보였으며, 특히 LightGBM은 높은 탐지율을 안정적으로 유지하면서도 학습·추론 시간이 짧아 실시간 대응이 요구되는 금융 운영 환경에 가장 적합한 것으로 확인되었다.

ACKNOWLEDGMENT

이 논문은 과학기술정보통신부 및 정보통신기획평가원의 정보통신방송 표준개발지원사업(RS-2024-00397768)사업으로 지원받은 연구결과임

참 고 문 헌

- [1] <https://www.kaggle.com/datasets/sgpjesus/bank-account-fraud-dataset-neurips-2022>.
- [2] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321?357, 2002, doi: 10.1613/JAIR.953.
- [3] P. Florek and A. Zagda?ski, "Benchmarking state-of-the-art gradient boosting algorithms for classification," *arXiv:2305.17094*, 2023, doi: 10.48550/arXiv.2305.17094.