모바일 금융 서비스 사기 거래 데이터를 활용한 이상 탐지 기법 비교

오재호[1], 김종휘[1], 박지영[1]
[1]광운대학교 정보융합학부

초 록

이상 탐지는 다양한 연구 및 응용 분야에서 활발히 연구되고 있는 문제이다. 또한, 본 논문에서 활용한 금융 사기 거래 데이터는 모바일 금융 서비스에서의 사기 및 이상 거래에 대한 문제를 담고 있는 데이터이다. 최근 이와 관련한 사회적 관심이 집중된 가운데, 기법적으로 이러한 문제를 풀기 위한 많은 노력들이 있다. 본 실험에서는 데이터의 기본적인 특징을 이해하고 정상과 비정상을 정의한 후, 각각의 이상탐지 알고리즘을 활용하여 해당 데이터의 정상을 탐지하는 학습모델을 구축하고자 한다. 특히, 수업을 통해 학습한 알고리즘 외에도 GANomaly 알고리즘을 스스로 학습하고 실험에 추가하였다. 실험에서는 GANomaly, Isolation Forest, PCA 등 7가지 알고리즘을 적용하여 성능 비교 실험을 수행했으며, 실험 결과 Isolation Forest가 F1-score 85.8%, Accuracy 86.3%으로 타 알고리즘에 비해 우수한 성능을 보였다. 본 실험을 통하여 이상탐지가 현장에 적용될 수 있음을 느꼈고, 실제 금융 사기 거래 탐지에 적용하여 사기 거래의 추가 피해를 막을 수 있는 수단으로써의 효과를 확인하였다.

1. 서론

정보통신기술의 발달로 모바일 금융 거래 서비스가 활성화되면서 모바일 송금 서비스를 활용한 거래가 활발히 이루어지고 있다.[1] 이에 금융 거래 내에서 사기 및 이상 거래에 대한 피해 또한 증가하고 있다.[2] 본 논문에서는 금융사기 및 이상 거래의 피해를 방지하기 위해 다양한 이상탐지 알고리즘을 비교 분석함으로써 실제 사기 거래에 대한 활용성을 입증하고자 한다.

본 논문의 구성으로 2장에서는 문제 상황과 데이터에 대해 소개하고, 3장에서는 탐색적 데이터 분석(EDA)를 통해 데이터의 분포 및 통계를 요약한다. 4장에서는 추가로 학습한 알고리즘을 소개하며, 5장 실험 결과 및 결론 부분에서는 연구 내용을 정리하며 마무리한다.

2. 문제 상황 및 데이터 설명

모바일 금융 서비스의 사기 거래 피해와 사회적 관심이 증가함에 따라 최근 이상 탐지에 대한 관심 또한 증가하고 있다.[3] 본 논문에서는 PaySim 시뮬레이터를 통해 인공적으로 생성된 모바일 금융 거래 데이터 세트를 이용한다. [4]의 연구에서는 PaySim 시뮬레이터는 실제 원본 데이터 세트를 기반으로 시뮬레이션 기술과 수학적 통계 기법을 이용하여 실제 데이터 세트만큼의 가치가 있다고 밝혔다. 이에 본 논문에서는 금융 사기 거래 탐지를 위한 이상 탐지 기법 비교 실험에 본 데이터를 사용하고자 한다.

본 데이터는 캐글(kaggle)에서 제공한 데이터로 원본 데이터의 변수는 다음과 같다. 30일 간의 시뮬레이션 기간 동안 단위 당 1시간을 의미하는 step, 거래 유형을 의미하는 type, 거래를 시작한 고객의 코드를 의미하는 nameOrig, 거래 전, 후 잔액을 의미하는 oldbalanceOrg, newbalanceOrig, 거래 수신자의 코드를 의미하는 nameDest, 수신자의 거래 전, 후의 잔액을 의미하는 oldbalanceDest, newbalanceDest, 대량 이체 시도 여부의 isFlaggedFraud, 사기 거래 여부이자 타겟 변수인 isFraud가 있다.

다음으로, 이상 탐지에 더 적절한 변수를 사용하기 위해 변수 가공(feature engineering)을 적용하였다. 먼저, 총 기간 동안의 시점을 나타내는 step변수는 시점이 지나면서 값이 커지게 되는데, 이에 대한 의미 보다는 하루 동안의 시간대를 나타내는 변수가 이상 탐지에 더 실용적이라 판단하였다. 따라서, step변수를 24로 나눈 나머지로 변경하여 하루 동안의 시간대를 나타내는 hours 변수로 변경하였다. 또한 거래 전, 후의 계좌 잔액의 크기 보다는 거래 전, 후의 계좌 잔액의 변화 값이 예측에 도움될거라 판단하여 ‘oldbalanceOrg’, ‘newbalanceOrig’, ‘oldbalanceDest’, ‘newbalanceDest’ 4가지 변수를 diff_balanceOrig, diff_balanceDest변수로 변경하여 사용하였다. 마지막으로 type 변수는 총 5가지의 카

테고리를 가진 변수로 레이블이 많지 않아, one-hot encoding을 적용하였다.

3. 탐색적 데이터 분석 (EDA)

본 장에서는 탐색적 데이터 분석(EDA)를 통해 실험에 사용한 데이터의 변수와 샘플의 수, 변수 별 분포와 결측치 및 이상치에 대해 설명 및 요약하도록 한다.

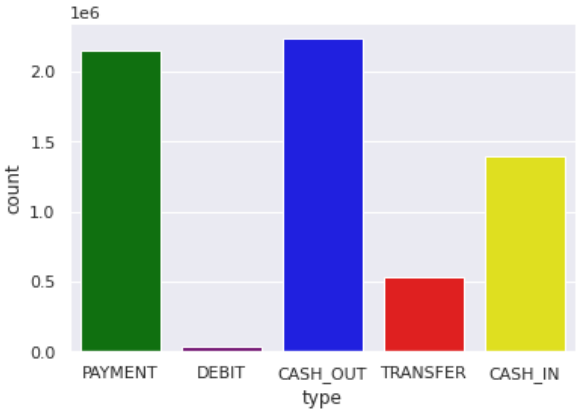
전체 데이터 세트와 훈련용, 평가용 데이터 세트의 범주 별 데이터 수는 다음 <표1>과 같다. 전체 데이터의 정상 데이터는 약 99.87%로, 비정상 데이터는 0.13%로 나타났으며, 범주 불균형 현상이 극심한 것으로 나타난다. 평가용 데이터는 비정상과 정상 데이터를 1:1 비율로 맞추어 구성하였고, 나머지 데이터는 훈련용 데이터로 이용하였다.

<표1> 데이터 세트의 범주 별 데이터 수

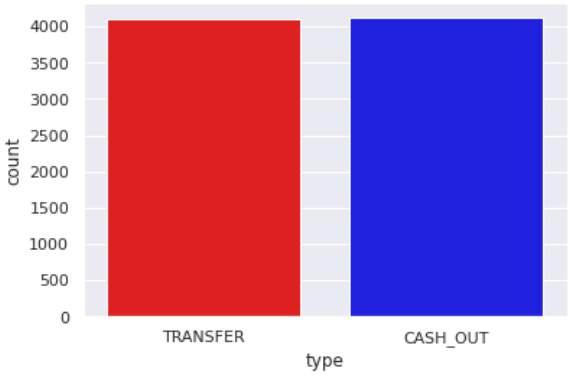
	전체 데이터	훈련용 데이터	평가용 데이터
정상	6,354,407	6,346,194	8,213
비정상	8,213	0	8,213
총합	6,362,620	6,346,194	16,426

이상 탐지 모델에 사용한 변수는 변수 가공을 거쳐 'amount', 'isFlaggedFraud', 'hours', 'diff_balanceOrig', 'diff_balanceDest', 'type' 6가지를 사용하였으며, type 변수는 one-hot encoding을 적용하였다. 또한, 모든 변수에 대해 결측치는 존재하지 않았다.

다음으로 각 변수별 데이터의 분포를 시각화하여 나타내었다. 사기 거래에 해당하는 비정상 범주의 데이터와 정상 범주 데이터의 분포가 뚜렷하게 차이나는 경우에 대해 분포 그래프를 첨부하였다.

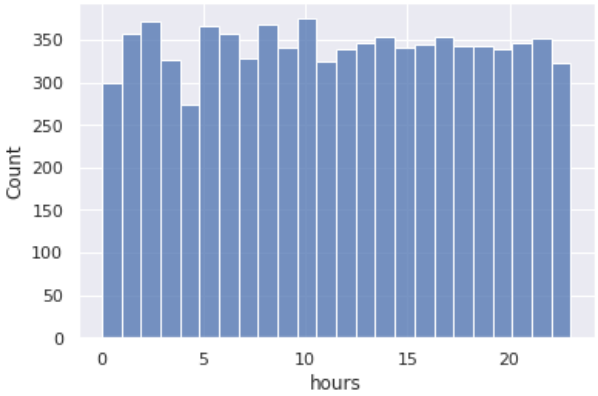


(그림1) 정상 데이터의 type 변수 범주 분포

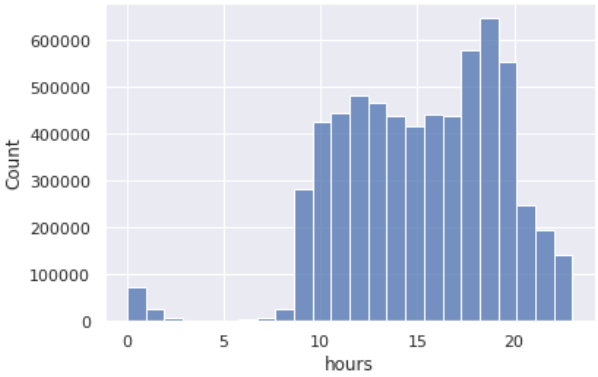


(그림2) 비정상 데이터의 type 변수 범주 분포

(그림1)과 (그림2)를 보았을 때, 비정상 범주 데이터는 type 변수에 대해 "TRANSFER", "CASH_OUT" 범주에 대해서만 나타났다.



(그림3) 정상 데이터의 hours 변수 분포



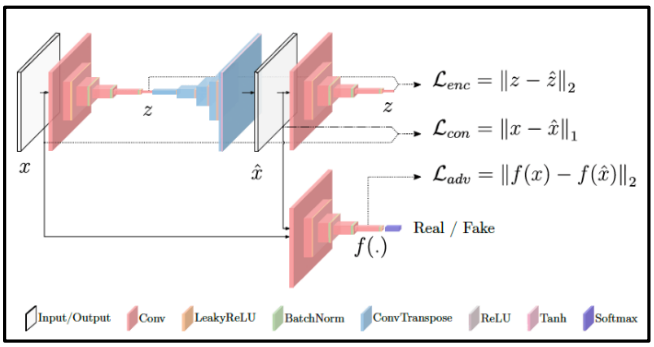
(그림4) 비정상 데이터의 hours 변수 분포

(그림3)과 (그림4)를 보았을 때는 hours 변수의 값이 다른 수치에 비해 0~8의 값에서 비정상 비율이 높은 것을 알 수 있다. 다른 변수에 대한 데이터 분포는 부록(Appendix)에 첨부하였다.

<표2> 이상탐지 알고리즘 성능 실험 결과

	Isolation Forest	LOF	PCA	Gaussian Density Estimation	Mixture of Gaussian	Auto Encoder	GANomaly
F1-score	0.858	0.636	0.828	0.634	0.857	0.783	0.815
Accuracy	0.863	0.879	0.839	0.705	0.873	0.804	0.83
Recall	0.827	0.737	0.778	0.511	0.759	0.709	0.898
Precision	0.892	0.56	0.885	0.834	0.984	0.875	0.745

4. GANomaly

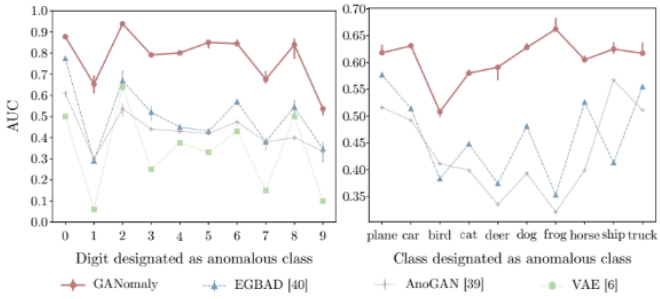


(그림5) GANomaly의 제안 방법

GAN을 사용하는 Anomaly Detection 방법론으로써, ACCV 2018에서 제안되었다.[5] GANomaly는 이미지를 생성하는 과정과 latent space를 학습하는 과정을 한 번에 진행하기 위해서 Generator와 Discriminator 이외에 인코더 부분이 추가되었다.

Adversarial Loss, Contexual Loss, Encoder Loss의 3가지 Loss Function을 사용하였으며, normal 이미지를 모델의 input으로 받아서 compress한 후에 다시 reconstructed하여 새로운 이미지를 생성한다.

Discriminator 모델은 기존과 동일하게 nomal 이미지와 생성된 이미지를 받아서 구분한다. 새롭게 추가된 인코더 모델은 생성된 이미지를 다시 compress하여 latent vector를 추출하는 방식으로 작동한다.



(그림6) MNIST와 CIFAR데이터 벤치마크 결과

(그림6)에서 GANomaly의 경우 이전의 모델(EDBAD, AnoGAN,VAE)에 비해 AUC가 높은 것으로 밝혀졌다.[5]

5. 실험 결과 및 결론

본 논문에서는 이상 탐지 알고리즘 별 성능 비교 실험을 위해 Isolation Forest, LOF(Local Outlier Factor), PCA(Pricipal Component Analysis), Gaussian Density Estimation, Mixture of Gaussian, Auto Encoder, GANomaly 7가지 모델을 사용하였고, <표1>의 훈련용 데이터와 평가용 데이터를 사용하여 평가를 진행하였다. 평가 지표로는 F1-score, Accuracy, Recall, Precision을 이용하였고, 결과는 <표2>와 같이 나타났다.

F1-score를 보았을 때, Isolation Forest 알고리즘이 약 85.8%로 가장 높은 성능을 보였고, 추가로 학습하여 실험했던 GANomaly는 약 81.5%을 기록하였다. Isolation Forest 알고리즘은 연산량이 적다는 특징을 가졌으며, 실제 실험결과에서도 약 7분 가량의 학습 시간이 소요되어 학습에 약 76분이 소요됐던 GANomaly에 비해 비용적, 성능적 측면에서도 우수한 것으로 나타났다. 이는 트리 구조를 베이스로 하여 정상 데이터만으로도 강건한 모델을 구성할 수 있다는 특징 때문으로 보인다.

또한 GANomaly의 경우, 전반적으로 우수한 성능을 보였다. 이상 탐지 벤치마크 데이터 세트에 대해 GANomaly가 Auto Encoder에 비해 더 우수한 성능을 보였던 이전 실험과 동일하게 본 논문에서 사용한 데이터 세트 또한 더 우수한 결과를 보였다. 하지만 타 알고리즘에 비해 연산량 및 학습 시간이 오래 걸린다는 단점 또한 존재했다.

본 논문을 통하여 이상탐지가 현장에 적용될 수 있을 만큼 높은 정확도를 보였고, 실제 금융 사기 거래 탐지 사례에 적용하여 금융 사기 거래의 추가 피해를 막을 수 있는 수단으로써의 효과를 확인하였다.

참고문헌

[1] Yoon, J., & Kim, B. (2022). 핀테크 확대가 금융안정에 미치는 영향 (Effects of Fintech Expansion on Financial Stability). Korea Deposit Insurance Corporation, 23(1-3).

[2] Park E.Y., Yoon J.W., “A Study of Accident

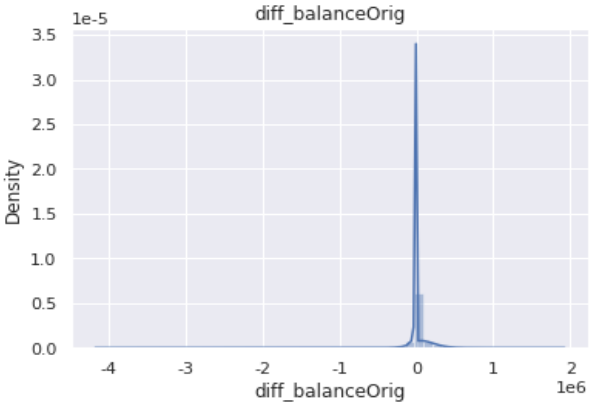
Prevention Effect through Anomaly Analysis in E-Banking", The Journal of Society for e-Business Studies, Vol.19, No.4, pp.119-134, November 2014

[3] 민기인, & 조위덕. (2021). 모바일 결제 환경에서의 실시간 이상거래 탐지를 위한 기계학습 알고리즘 성능비교 연구. 한국통신학회 학술대회논문집, 164-165.

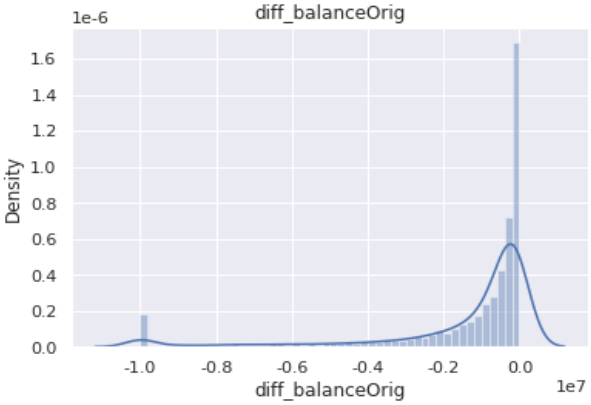
[4] Edgar Lopez-Rojas, Ahmad Elmir, and Stefan Axelsson. Paysim: A financial mobile money simulator for fraud detection. In 28th European Modeling and Simulation Symposium, 2016.

[5] Samet Akcay, Amir Atapour-Abarghouei, and Toby P. Breckon. 2018. GANomaly: Semi-supervised anomaly detection via adversarial training. In ACCV. Springer, 622–637.

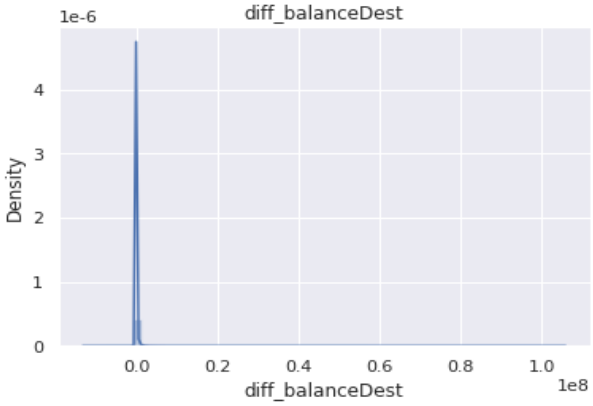
부록 (Appendix)



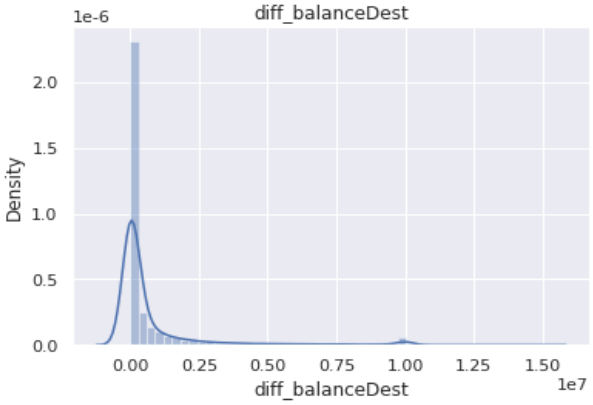
(그림7) 정상 데이터의 diff_balanceOrig변수 분포



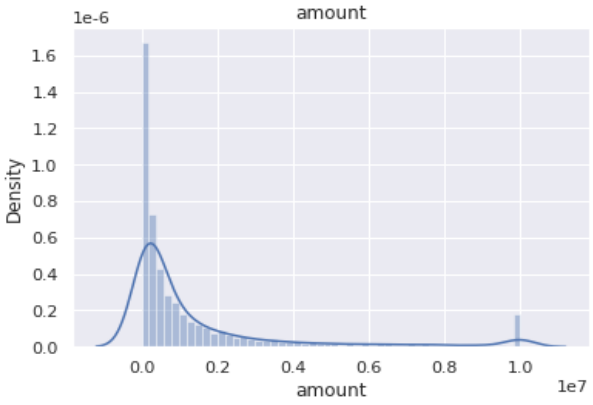
(그림8) 비정상 데이터의 diff_balanceOrig변수 분포



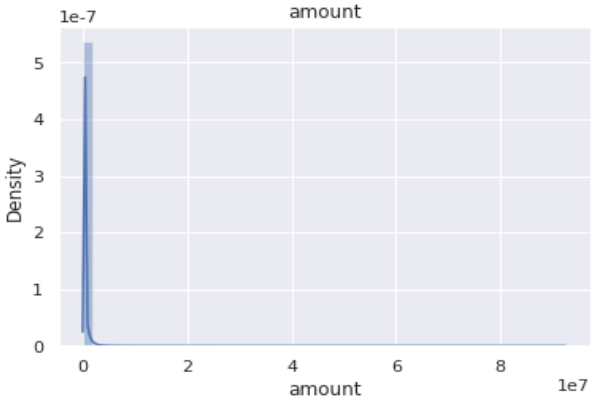
(그림9) 정상 데이터의 diff_balanceDest 변수 분포



(그림10) 비정상 데이터의 diff_balanceDest 변수 분포



(그림11) 정상 데이터의 amount 변수 분포



(그림12) 비정상 데이터의 amount 변수 분포