

# 시간대를 고려한 신용카드 이상탐지 기법 연구

A Study on Credit Card Anomaly Detection Considering Time Sequences

숙명여자대학교 빅데이터분석융합학(협동과정)

양 소 연



# 목차

1. 연구 배경 및 목적
2. 관련 연구
3. 시스템 모델
4. LSTM Autoencoder
5. 결론 및 향후 계획


# 1. 연구 배경 및 목적



핀테크(FinTech)

보안상의 취약점을 노린 금융사기 급증



 [금융위원회] 가족과 지인을 사칭하여 카카오톡, 휴대전화문자 등을 통해 돈과 개인정보를 요구하는 메신저피싱이 증가하고 있으니, 유의하시기 바랍니다.

[금융위원회] (대처방법)가족, 지인 여부를 통화 등을 통해 확인! 출처 불분명한 앱설치 거절! 신고는 금감원 ☎1332, 경찰청 ☎112

오후 2:01

# 1. 연구 배경 및 목적

## 이상금융거래 탐지시스템 (FDS, Fraud Detection System)



“분석 및 탐지” 기능에 대한 실험 수행

# 1. 연구 배경 및 목적

## 연구 문제 정의 및 해결방안

### 데이터 라벨링(Labeling)

- Labeled data 확보가 어려움
- 혼재되어 있음



- Labeled data를 활용한 지도학습
- Unlabeled data를 활용한 비지도 학습



- LSTM Autoencoder를 사용한 비지도학습

### 데이터 불균형

- FDS는 이진 분류 예측 모델
- 정상 / 이상 비율의 차이가 큼



- 오버샘플링 (Over-sampling)
- 언더샘플링 (Under-sampling)



- CTGAN을 활용한 합성 데이터 생성

## 2. 관련 연구

### 데이터 라벨링(Labeling)

- 이상 데이터와 레이블 데이터 확보의 어려움
- 이상 패턴의 기준이 바뀌면 새로운 레이블 데이터 생성해야 함

---

### • 지도학습

- LSTM과 Adaboost를 활용하여 0.996%의 민감도를 얻음(Esenogho et al, 2022)
- Random Forest를 활용하여 98.6%의 accuracy를 얻음(Gaikwad et al, 2022)

### • 비지도학습

- Autoencoder를 사용하여 변수를 추출하고 SVM 모델과 결합(Deng et al, 2020)

## 2. 관련 연구

### 데이터 불균형

- 리샘플링(resampling)을 활용하여 클래스 분포를 재조정
- 
- **SMOTE**(Synthetic Minority Over-sampling Technique)
  - SMOTE를 활용한 리샘플링 데이터로 0.9262의 accuracy를 얻음 (Ishaq et al, 2021)
- 
- **GAN**(Generative Adversarial Networks)
  - GAN을 활용한 리샘플링 데이터로 recall 점수를 5% 상승시킴 (Ngwenduna et al, 2021)

### 3. 시스템 모델

#### 데이터 설명

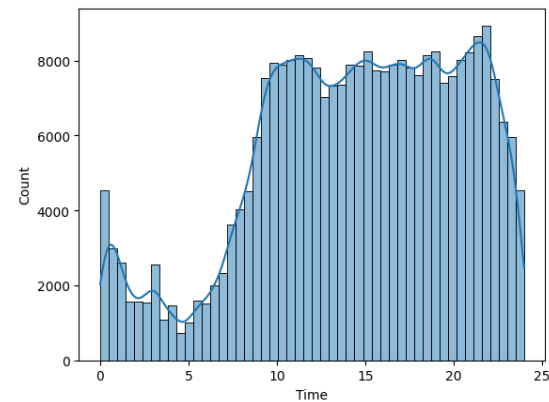
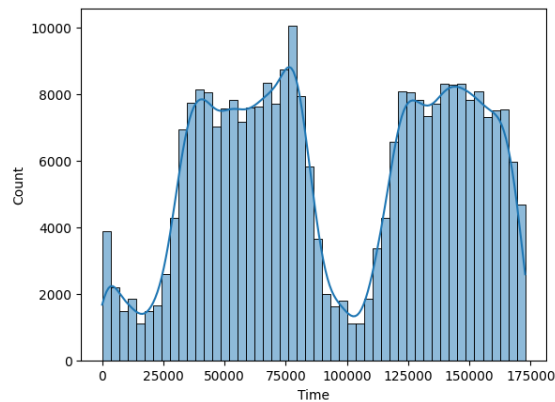
- Kaggle, "Credit Card Fraud Detection" (<https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud>)
  - 2013년 유럽에서 사용된 신용카드 거래 데이터
  - 약 28만개 데이터 중 492개만이 사기 데이터 → 0.172%
- 주요 변수는 V1, V2, ... V28 과 'Time', 'Amount' 그리고 'Class'
  - Time : 첫번째 거래와 마지막 거래 사이의 경과 시간(초)
  - Amount : 거래 금액, 비용을 고려한 학습에 사용 가능
  - Class : 사기 데이터는 1, 그렇지 않으면 0으로 표기
- 주성분 분석(PCA, Principal Component Analysis)을 통한 비식별화



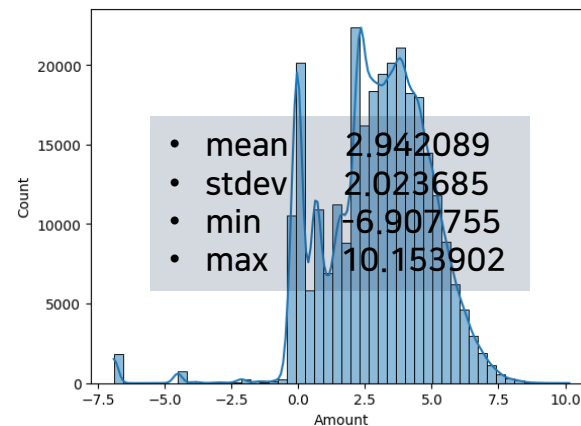
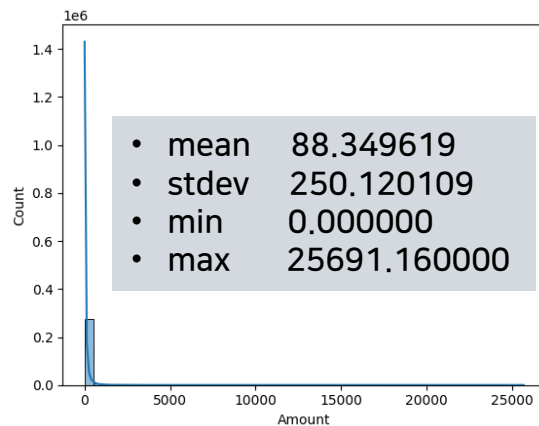
# 3. 시스템 모델

## 데이터 전처리

- Time을 일중 시간으로 변환



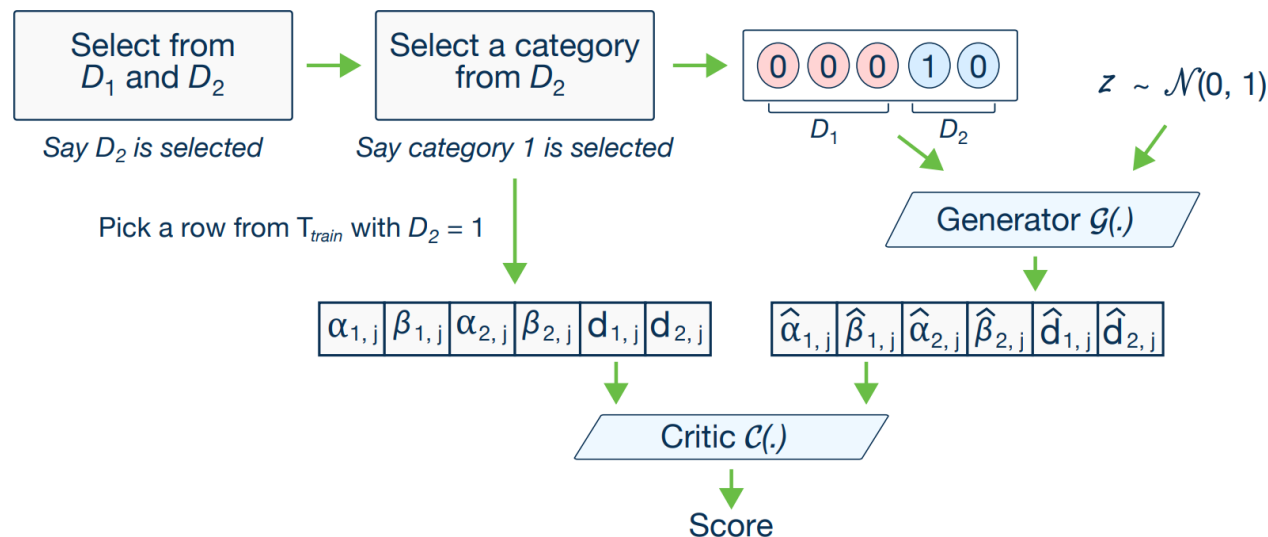
- Amount의 편차가 크므로 log를 이용하여 변환



# 3. 시스템 모델

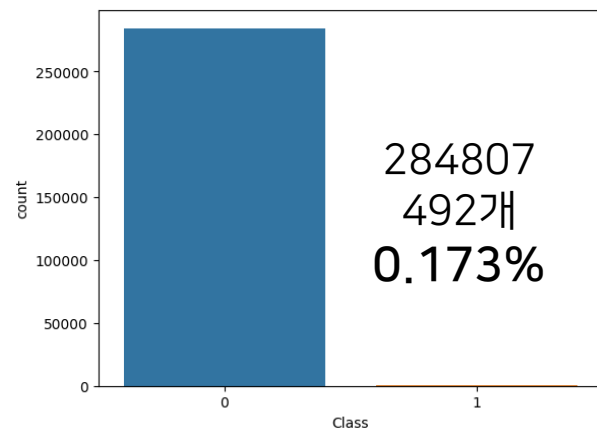
## CTGAN(Conditional Tabular Generative Adversarial Networks)

- 일반적으로 정형데이터는 연속형 변수와 카테고리 변수가 섞여 있음
- 기존의 tabular data GAN은 metric 문제가 발생
- 이러한 문제를 해결하기 위해 조건부 생성기를 사용하는 CTGAN을 설계

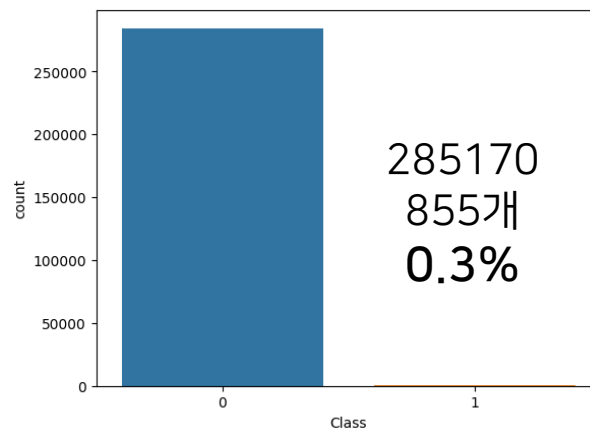


### 3. 시스템 모델

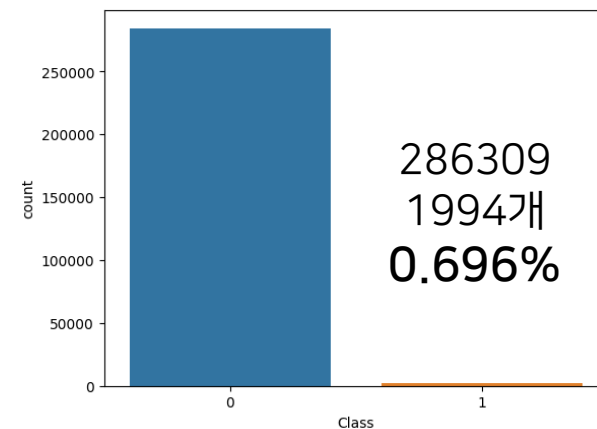
#### CTGAN을 활용한 합성 데이터 생성



Shared data



Fraud 0.3%



Fraud 0.7%

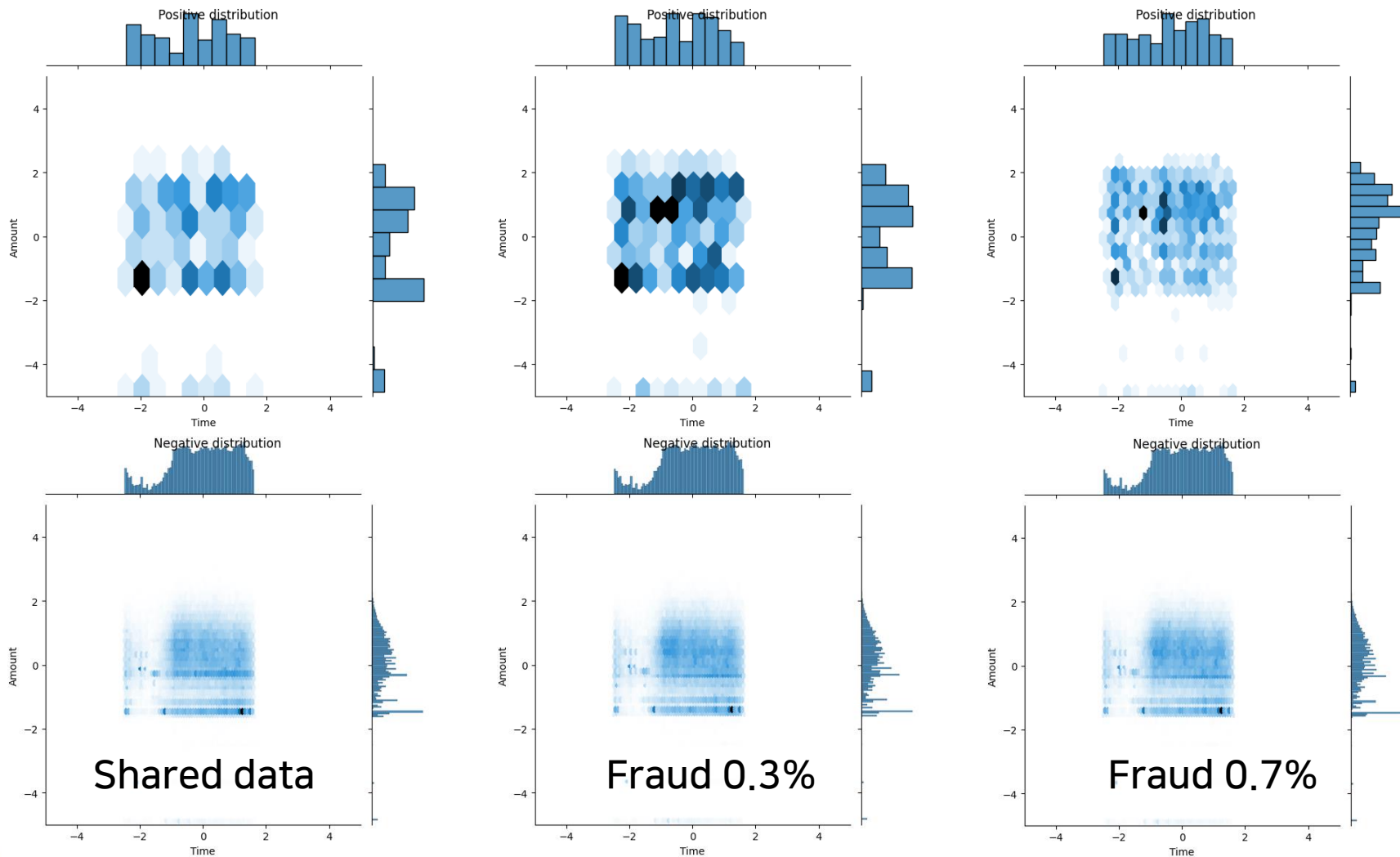
#### GAN 알고리즘 중

Tabular data에 최적화 되어있는 CTGAN을 활용하여

소수의 이상 데이터를 오버샘플링

# 3. 시스템 모델

## CTGAN을 활용한 합성 데이터 분포



### 3. 시스템 모델

#### 변수 선택(Feature Selection)

- 모델을 개발할 때 입력 변수를 줄이는 과정
- 입력 데이터의 개수를 줄여 모델링의 계산 비용을 절감
- 모델의 성능을 개선
- 일반적으로 통계적 기법을 사용 → XAI 기법 활용
- 입력 데이터와 목적 변수의 상관관계를 살펴본 후 선택

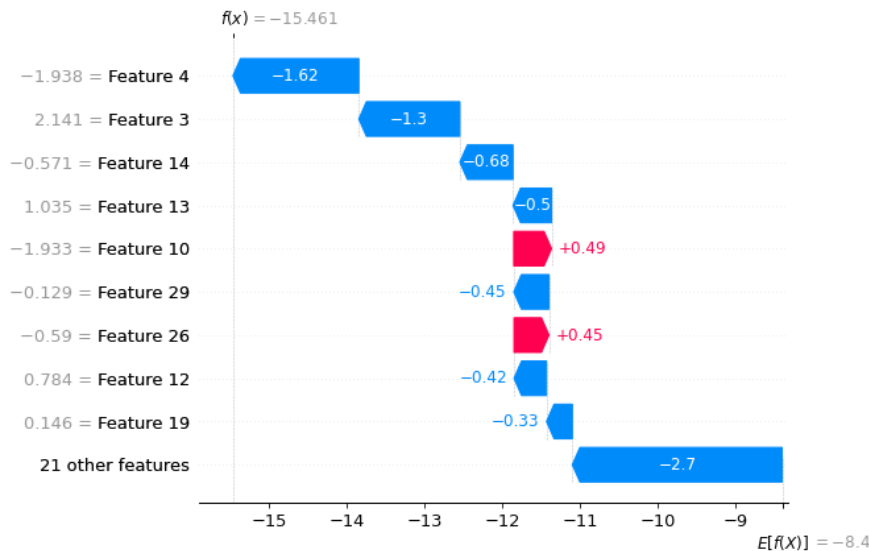
# 3. 시스템 모델

## XAI(eXplainable Artificial Intelligence)

AI의 블랙박스 문제를 해결하고, 인간에게 AI의 행동을 이해하기 쉽게 설명해주는 것이 목적

### SHAP

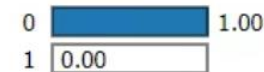
(Shapley Additive exPlanations)



### LIME

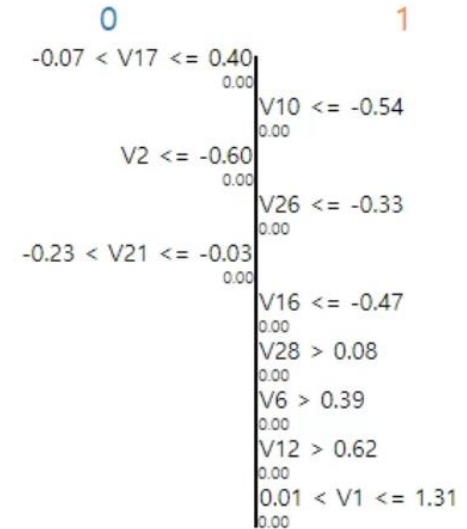
(Local Interpretable Model-agnostic Explanation)

Prediction probabilities



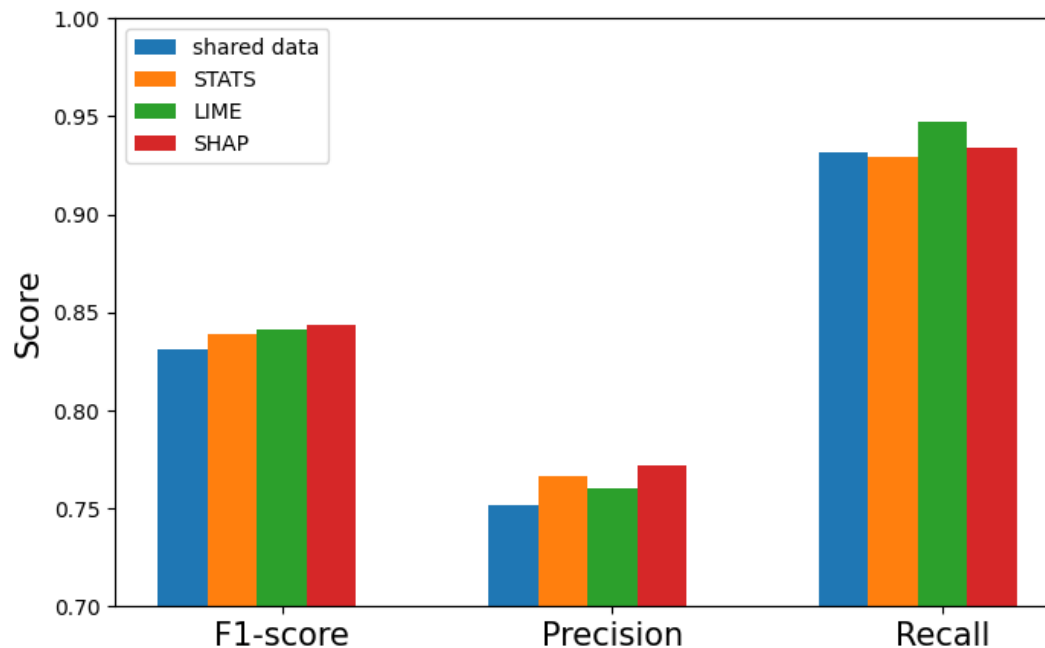
Feature Value

V17	0.13
V10	-0.57
V2	-2.74
V26	-0.75
V21	-0.21
V16	-2.14
V28	0.10
V6	1.02
V12	0.98
V1	0.34



### 3. 시스템 모델

#### Feature Selection 기법에 따른 성능 비교

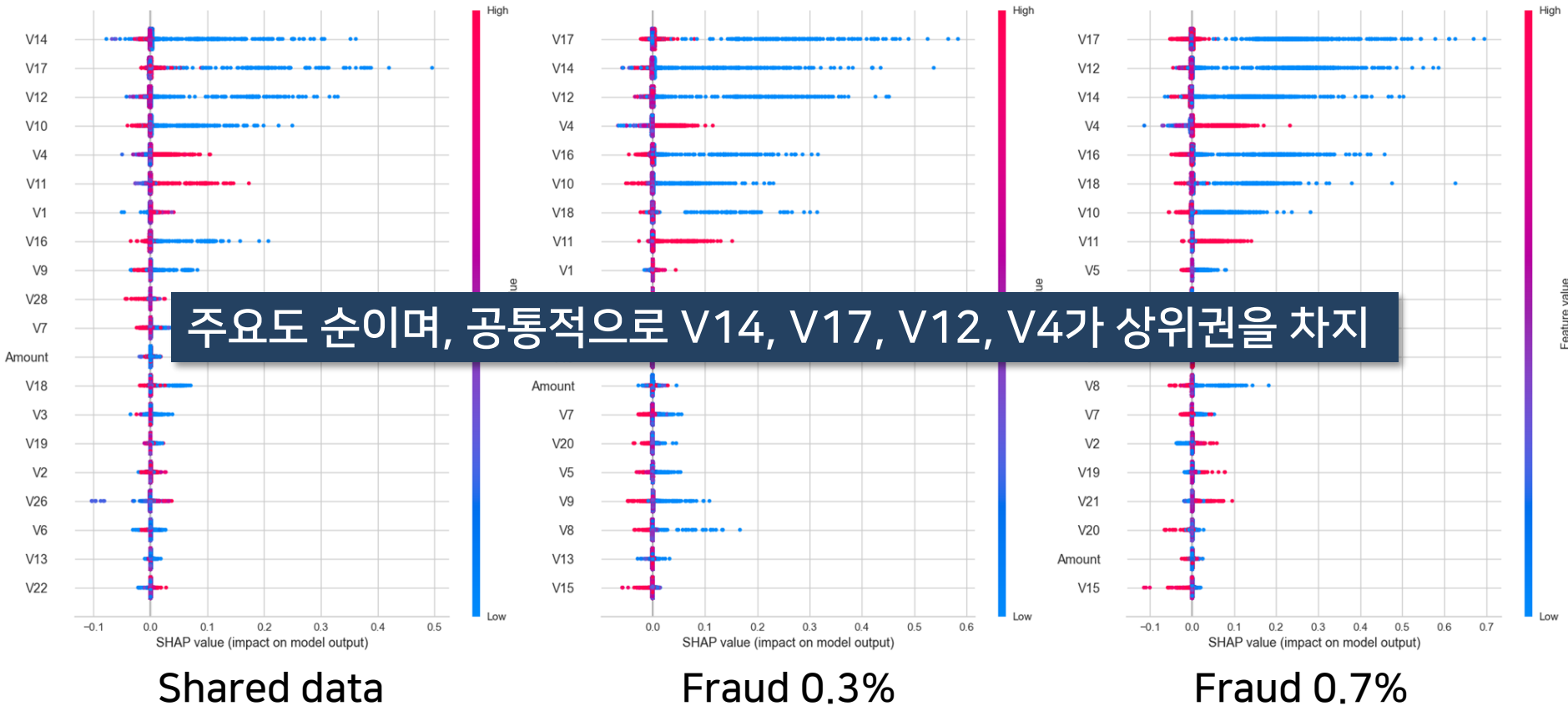


#### Random Forest를 사용하여 비교 분석

- 통계적 기법(STATS) : 피어슨 상관 계수, 카이제곱 검정, 차이분석, 회귀분석을 이용하여 선별
- LIME : 국소 회귀의 적합성만을 고려
- SHAP : 전체 데이터 샘플을 활용

# 3. 시스템 모델

## SHAP를 활용하여 추출한 주요 속성값





# 3. 시스템 모델

## 평가 방식 (Metric)

		실제 정답	
		True	False
분류 결과	True	True Positive (정답)	False Positive (오답)
	False	False Negative (오답)	True Negative (정답)

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Precision과 Recall의 조화평균

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$



# 제안 알고리즘

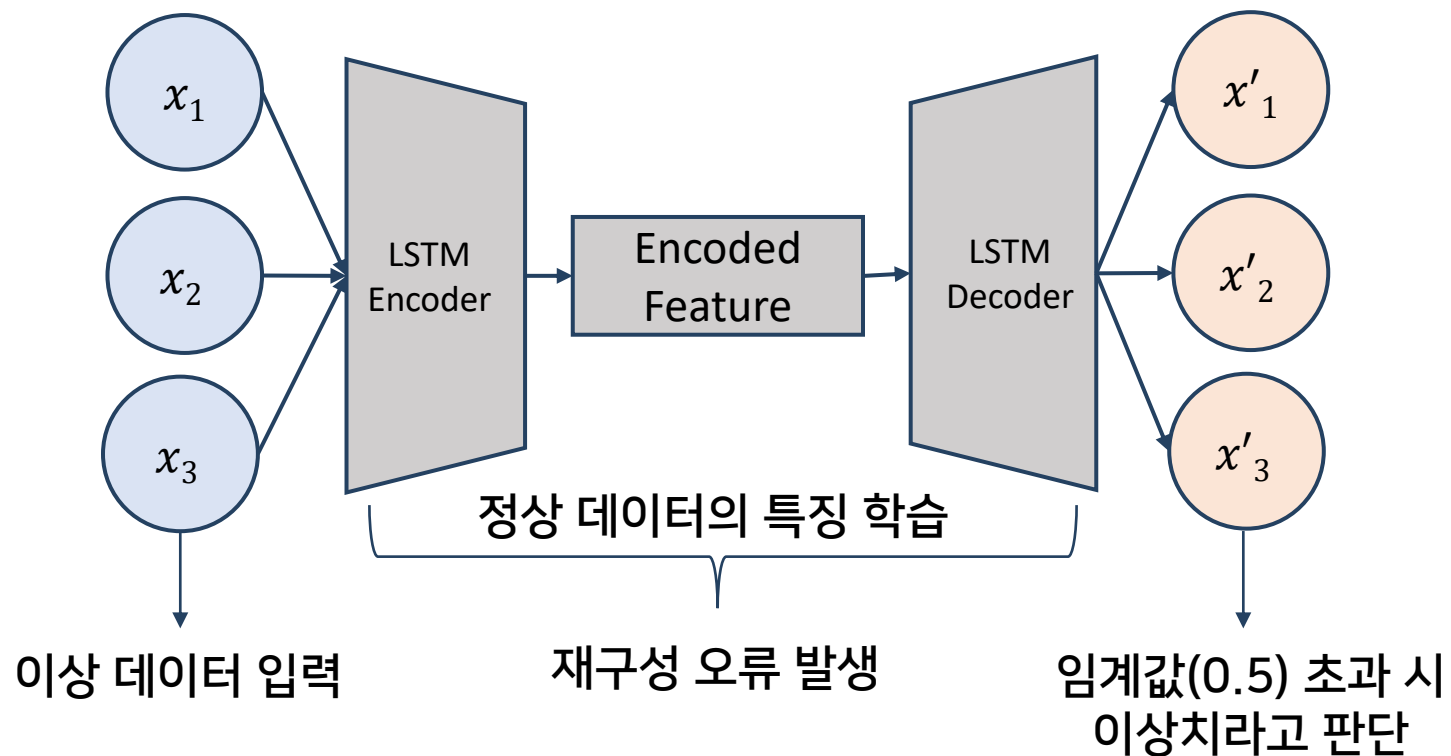
## 데이터 라벨링

- 현실적으로 라벨링 된 데이터가 부족 → Autoencoder 기반의 모델 사용
- 과거의 패턴도 일부 반영하여 참고하기 위해 시계열 기반의 LSTM 모델 활용
  - Autoencoder에 정상 데이터만 학습시켜 특징을 기억하도록 함 → LSTM의 메모리 활용
  - 해당 정상치 범위에서 벗어나는 데이터를 이상치로 감지

## 데이터 불균형

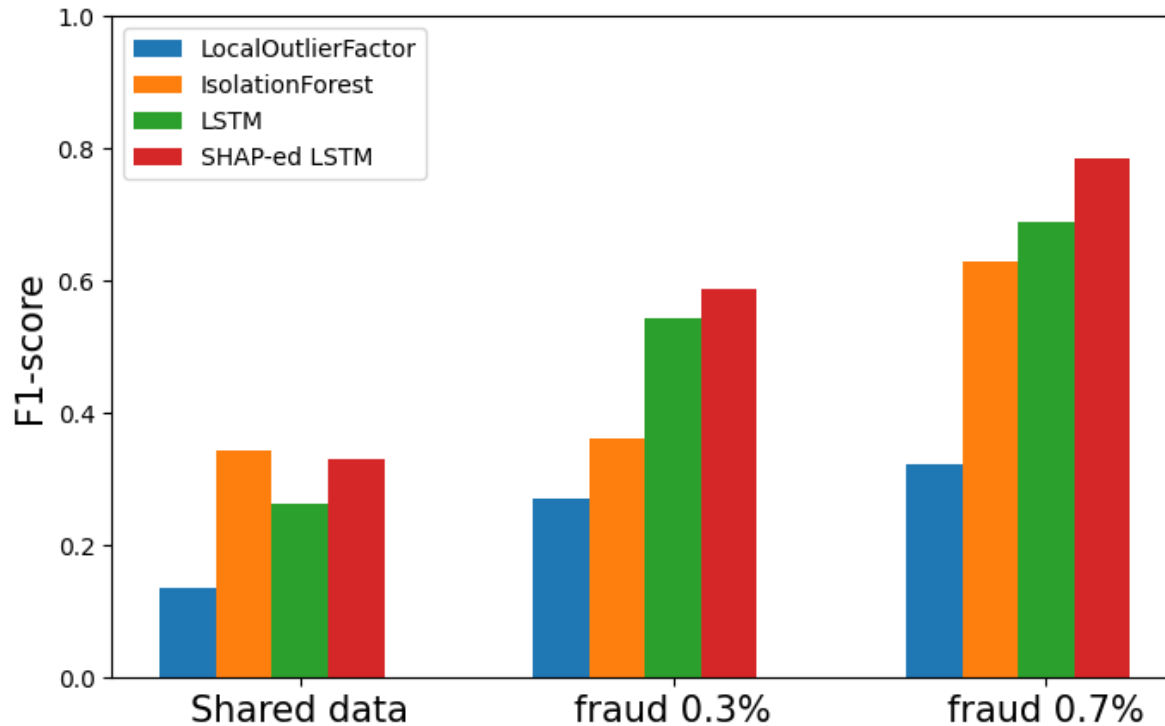
- 보통의 기계학습 모델은 클래스 비율이 동일하다고 가정
- 이상 탐지 모델 특성에 따라 Conditional Tabular GAN을 활용
- 이상 데이터의 비율을 각각 0.3%, 0.7%으로 리샘플링

## 4. LSTM Autoencoder



## 4. LSTM Autoencoder

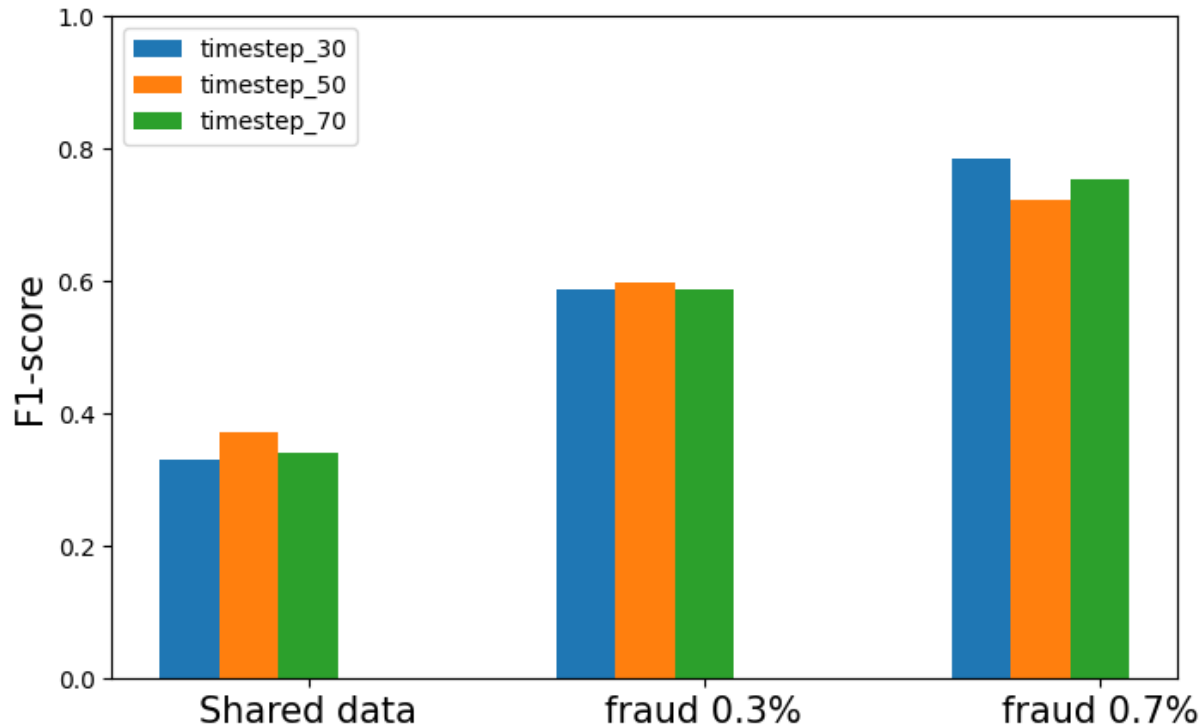
### 전통적 비지도학습 기법과의 비교



- Shared data에서는 Isolation Forest 와 제안 알고리즘이 비슷한 성능
- 이상치 비율이 높아질수록 제안 알고리즘인 SHAP-ed LSTM 의 성능이 우수

## 4. LSTM Autoencoder

### 최근 데이터 반영 개수에 따른 모델 성능 비교



이상치 비율이 낮으면 데이터의 전반적인 흐름을 보아야 하지만  
이상치 비율이 높으면 최근 데이터만 보는 것이 이상금융거래탐지에 유리

## 5. 결론 및 향후 계획

### 데이터 라벨링

- Autoencoder 모델을 기반으로 데이터 라벨링 문제 완화

### 데이터 불균형

- Conditional Tabular GAN을 이용하여 리샘플링

### 컴퓨팅 파워 절약

- XAI 기법을 활용하여 Feature 추출

### 최근 데이터 반영

- 사기 데이터 비율이 낮을 때는 전반적인 추이 분석
- 사기 데이터 비율이 높을 때는 최근 데이터에 집중

### 모델 해석

- XAI 기법의 SHAP 사용

### 향후 계획

- 최근 데이터를 얼마나 반영해야 하는지 optimization 연구 진행