**신용카드 사기 검출**

**신용카드 사기 검출**

**프로젝트 개요 및 목표**

본 프로젝트는 실제 신용카드 거래 데이터를 기반으로, 사기 거래를 탐지하는 머신러닝 분류 모델을 개발하는 데 목적이 있습니다. 극심한 클래스 불균형 문제를 해결하기 위해 다양한 전처리, 불균형 처리(SMOTE, 클래스 가중치), 하이퍼파라미터 튜닝을 적용하여 정확도뿐 아니라 Precision, Recall, F1 Score, AUC 등을 종합적으로 고려한 실용적 모델을 도출하였습니다.

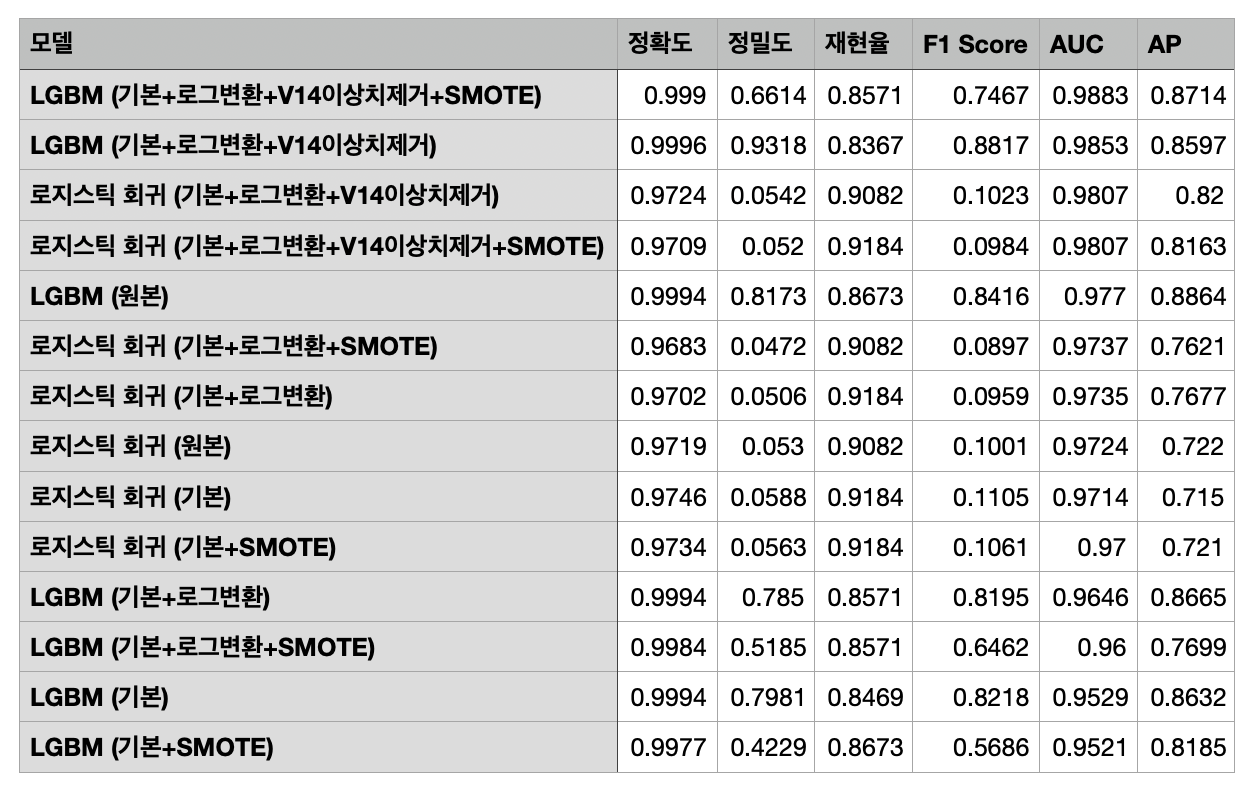
**데이터셋 요약**

• 총 거래 수: 284,807건  
• 타겟 변수: Class (0=정상, 1=사기)  
• 클래스 비율: 정상 99.83%, 사기 0.17% (극단적 불균형)  
• 피처 수: 30개 (Time, Amount, V1~V28 PCA 기반 특성)

**주요 전처리 전략**

• Amount 로그 변환 및 표준화  
• Time 제거  
• V14 이상치 제거 (정상 거래 중 약 1,100건)  
• SMOTE는 훈련셋에만 적용하여 소수 클래스 보정

**모델 비교 결과 요약**



• LightGBM은 모든 조합에서 로지스틱 회귀보다 Precision, F1 Score가 우수  
• 최종 조합(LGBM + 로그변환 + 이상치 제거)의 F1 Score는 0.86, Precision 86.5%, Recall 84.7%, AUC 0.989  
• SMOTE는 Recall은 향상되지만 Precision이 크게 떨어져 실무 적용에 주의 필요

**최종 모델 요약**

• 선정 모델: LightGBM  
• 적용 전처리: Time 제거, Amount 로그 변환+정규화, V14 이상치 제거  
• 성능 요약: Precision 86.5%, Recall 84.7%, F1 Score 0.8557, AUC 0.9892

**실무 적용 인사이트**

• Accuracy만으로 평가 불가. Precision, Recall, F1, AUC 등 다양한 지표 고려 필요  
• 임계값 조정 및 정책 기반 튜닝이 실무 적용에 중요  
• LightGBM은 예측 속도, 해석력, 실시간 적용성이 뛰어남  
• 간단한 전처리만으로도 성능 20~30% 이상 개선 가능

1. 프로젝트 개요 및 목표

# 1.1 프로젝트 개요

신용카드 사기 탐지(Credit Card Fraud Detection)는 과거의 정상 거래와 사기 거래 데이터를 학습하여 새로운 거래가 발생했을 때 이 거래가 사기일 가능성이 있는지를 자동으로 판별하는 분류 문제입니다.

이번 프로젝트의 목표는 신용카드 거래 데이터를 기반으로 머신러닝 분류 모델을 구축하고 다양한 전처리 전략 및 불균형 처리 기법 그리고 하이퍼파라미터 튜닝을 통해 성능을 개선하며 최종적으로 실무에 적용 가능한 탐지 모델을 도출하는 것입니다.

# 1.2 데이터셋 소개

본 프로젝트에서 사용한 데이터는 유럽 내 실제 신용카드 거래 데이터를 기반으로 구성된 공개 신용카드 사기 탐지 데이터셋입니다.

* 총 거래 수 : 284807건
* 피처 수 : 30개(V1 ~ V28은 PCA 변환된 특성 + Time, Amount)
* 타겟 변수 : Class(0 = 정상거래, 1 = 사기거래)
* 클래스 비율 : 정상 거래 약 99.83%, 사기거래 약 0.17% → 극심한 클래스 불균형

# 1.3 문제 정의 및 목표

이 프로젝트는 이진 분류 문제로 정의됩니다.  
즉, 주어진 거래 데이터의 피처들을 바탕으로 해당 거래가 사기(Fraud)인지 아닌지(Non-Fraud)인지를 예측하는 것이 목적입니다.

모델의 성능을 평가하기 위해 다음과 같은 종합적인 분류 지표를 사용하였습니다.

* Accuracy(정확도) : 전체 예측 중 정답 비율. 불균형 데이터에서는 오해를 줄 수 있으므로 참고용으로 사용.
* Precision(정밀도) : 사기라고 예측한 것 중 실제 사기 거래 비율.

→ False Positive(정상 거래를 사기로 오탐) 방지에 중요.

* Recall(재현율) : 실제 사기 거래 중 모델이 탐지한 비율.

→ False Negative(사기를 놓치는 경우) 방지에 중요.

* F1 Score : 정밀도와 재현율의 조화 평균. Precision과 Recall 간 균형 평가에 유용.
* ROC AUC : 다양한 임계값에서의 전체 분류 성능을 종합적으로 평가.
* AP(Average Precision) : Precision-Recall 곡선 면적. 불균형 데이터에서 유용한 종합 지표.

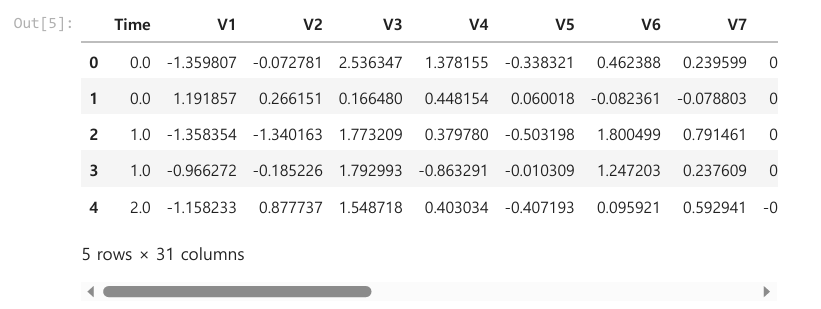
본 프로젝트에서는 정확도(Accuracy)만으로는 모델 성능을 판단하지 않고

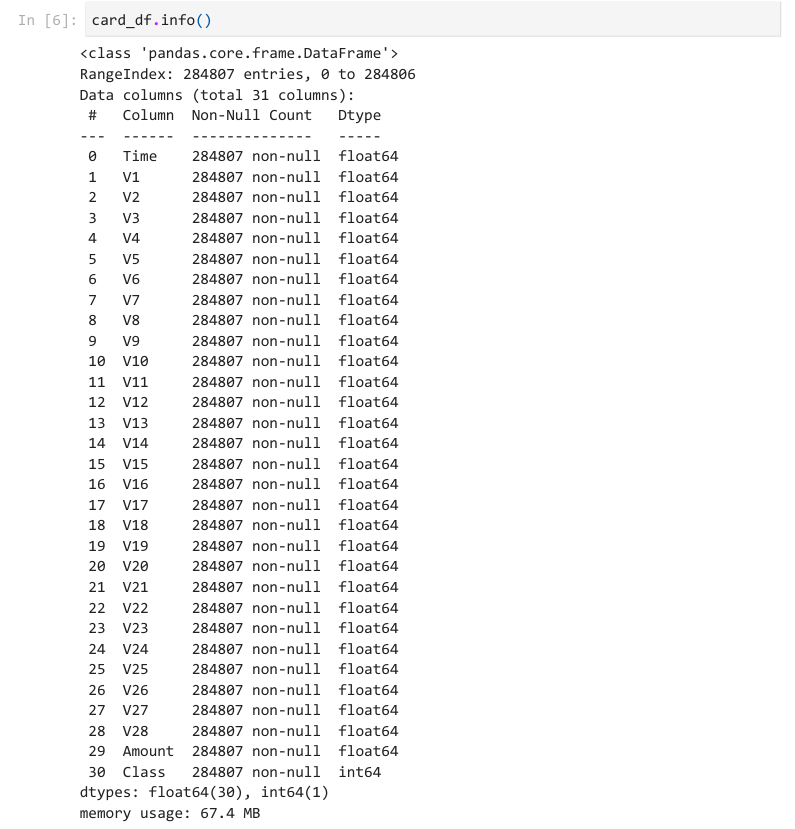
F1 Score, Precision, Recall, AUC, AP를 함께 고려해 보다 신뢰성 있고 현업에 적용 가능한 모델을 선정하였습니다.

1. 데이터 분석(EDA)











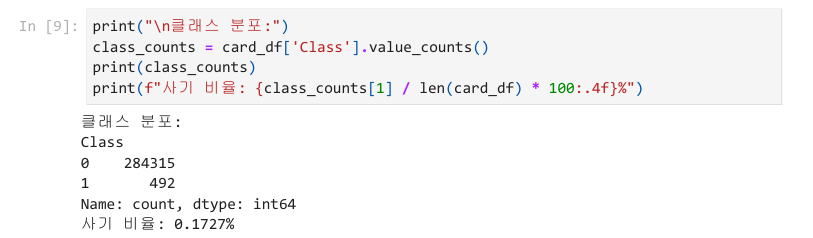


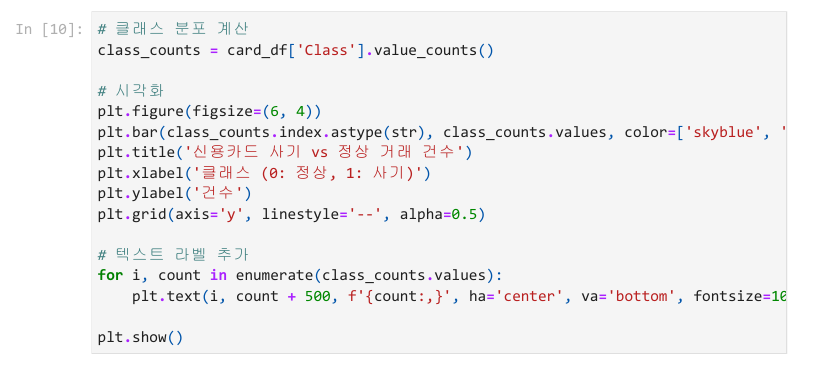
# 2.1 클래스 불균형 확인

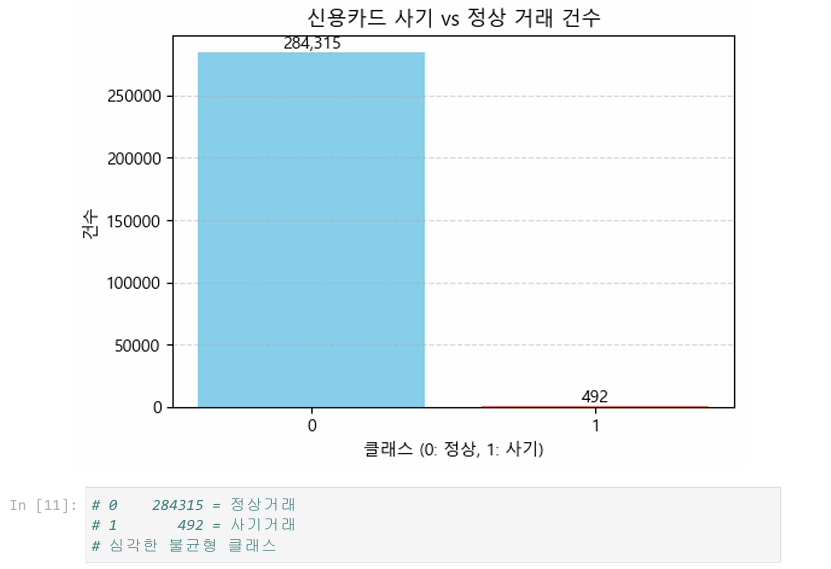
전체 284,807건의 거래 중 사기 거래는 단 492건(약 0.17%)에 불과하며 나머지 99.83%는 정상 거래로 구성되어 있습니다.

→ 이는 극단적인 클래스 불균형 상황으로 모델이 대부분을 정상 거래로 예측해도 Accuracy는 매우 높게 나올 수 있습니다.

따라서 Accuracy보다는 Precision, Recall, F1 Score, AUC 등의 정교한 평가 지표가 중요하며 데이터 전처리 및 학습 시 불균형 처리(SMOTE, 클래스 가중치 등)가 핵심 이슈가 됩니다.







# 2.2 거래 금액(Amount) 분포

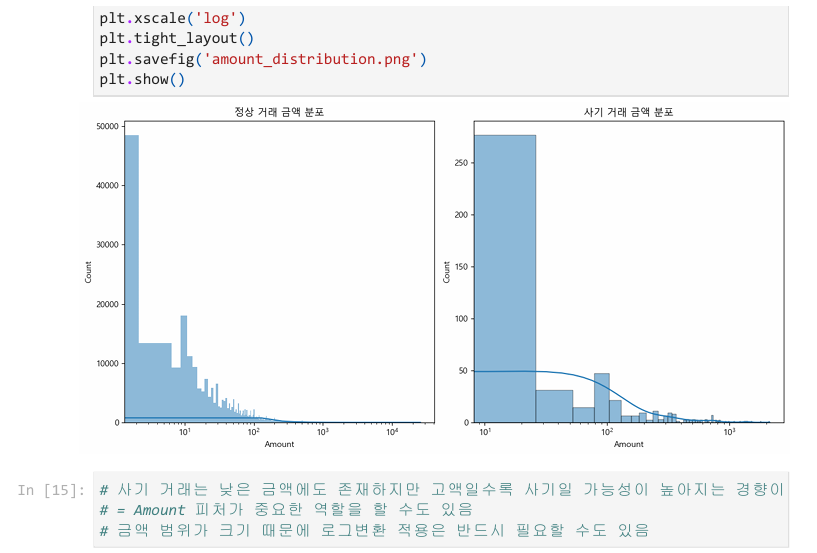
Amount 피처는 0에서 25,691 유로까지의 분포를 가지며 대부분은 소액 거래지만 일부는 수천~수만 유로에 달하는 거래도 존재합니다.

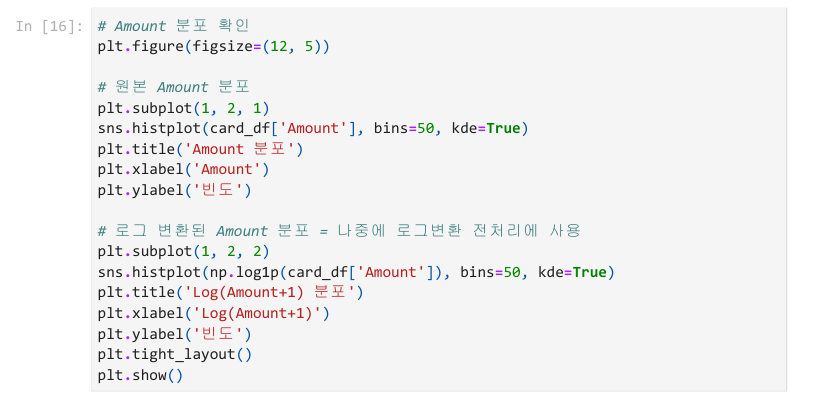
→ 전체적으로 오른쪽 꼬리가 긴(right-skewed) 분포 형태이며 로그 변환(log1p)을 통해 정규성에 더 가깝게 변환할 수 있습니다.

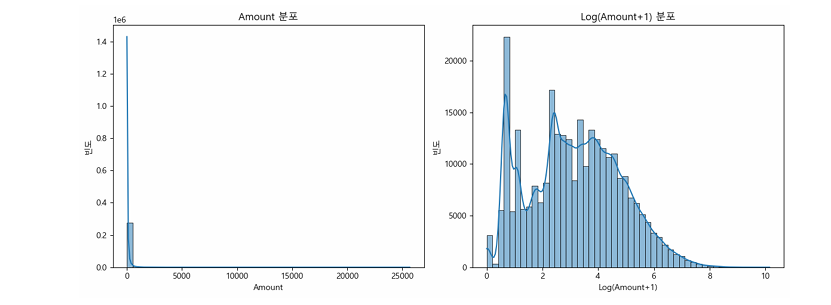
사기 거래는 고액이든 소액이든 다양한 금액대에 분포되어 있어 단순 금액 기준 탐지는 어렵습니다.

따라서 Amount 피처는 로그 변환 + 정규화(StandardScaler)를 통해 스케일 조정이 필요합니다.









# 2.3 시간(Time) 피처

Time 피처는 기준 시점부터의 경과 시간을 초 단위로 나타내며, 약 172,800초(48시간) 동안의 거래를 포함합니다.

→ 사기 거래가 특정 시간대에 몰려 있는 경향은 뚜렷하지 않으며 모델 성능에 크게 기여하지 않는 것으로 판단되어 전처리 과정에서 제거하였습니다.

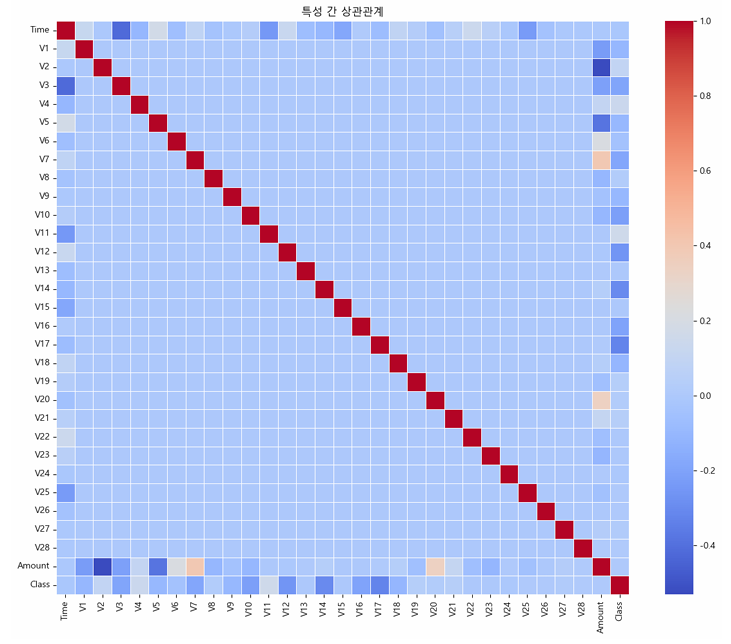
# 2.4 피처 분포 및 상관관계

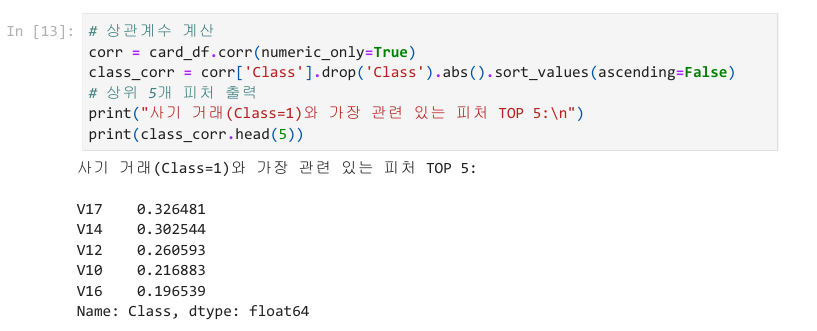
익명화된 PCA 피처(V1 ~ V28)는 대부분 평균 0에 가까운 정규 분포 형태를 보이며 일부 피처에서는 정상 거래와 사기 거래 간 뚜렷한 분포 차이가 관찰됩니다.

* 예 : V4, V11, V12, V14, V18 등은 두 클래스 간 값 분포가 눈에 띄게 다름
* 반면, V25, V26, V28 등은 분포 차이가 적어 예측력 낮음

일부 피처는 사기 거래에 대한 판별력이 높아 모델에서 중요한 역할을 할 것으로 기대됩니다.







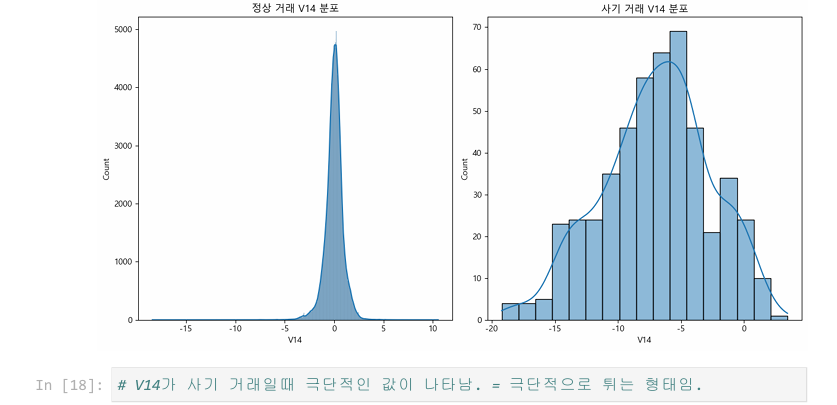
2.5 이상치 탐지

특히 V14 피처에서 정상 거래임에도 사기 거래와 유사한 극단적 값(-18 이하)을 가진 샘플이 존재했습니다.

→ 이는 모델이 혼란스러워할 수 있는 이상치로 작용할 수 있으며 정상 거래 중 이상치 샘플은 제거하여 모델 학습 안정성 향상을 기대할 수 있습니다.

전처리 과정에서 V14 기준으로 IQR 방식 이상치 제거를 적용하여 정상 거래 중 약 1,100건의 이상치를 제거하였습니다.





1. 데이터 전처리

EDA 결과를 바탕으로 몇 가지 데이터 전처리 전략을 수립하고 적용하였습니다.  
전처리는 누적 적용 방식으로 구성되었으며 총 4가지 전처리 조합을 실험하고 비교 평가하였습니다.

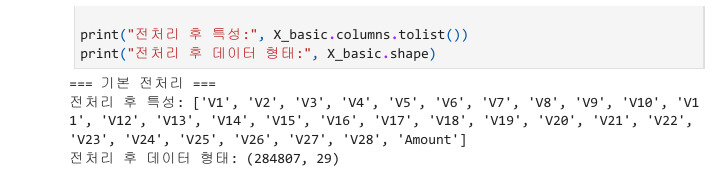
# 전처리 없음(원본 데이터)

* Time, Amount, V1~V28 피처를 그대로 사용
* Time 피처는 시계열 정보로 해석하기 어려우며 유의미한 정보가 부족하다고 판단되지만 비교를 위한 baseline 조합으로 포함

# 기본 전처리 : Time 피처 제거 + Amount 표준화

* Time은 사기 여부와 직접적인 상관관계가 약하고 순차적 패턴 분석이 목적이 아니기 때문에 제거
* Amount는 금액 크기 차이로 인해 모델 학습 시 영향력이 과도하게 커지는 현상 방지를 위해 z-score 방식으로 정규화

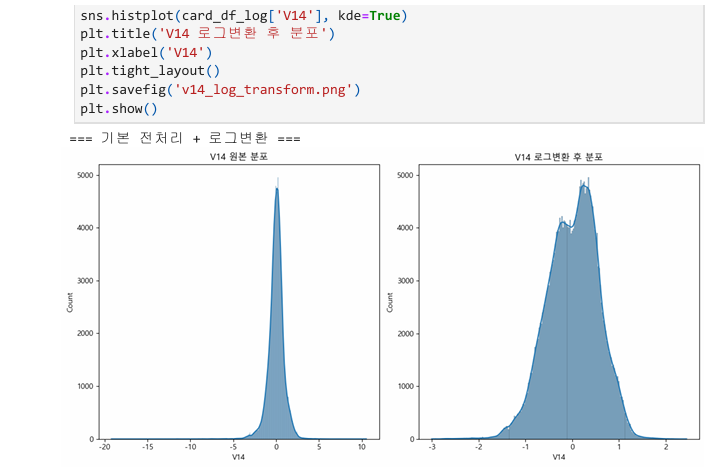




# 로그변환 추가 : 기본 전처리 + Amount 로그 변환

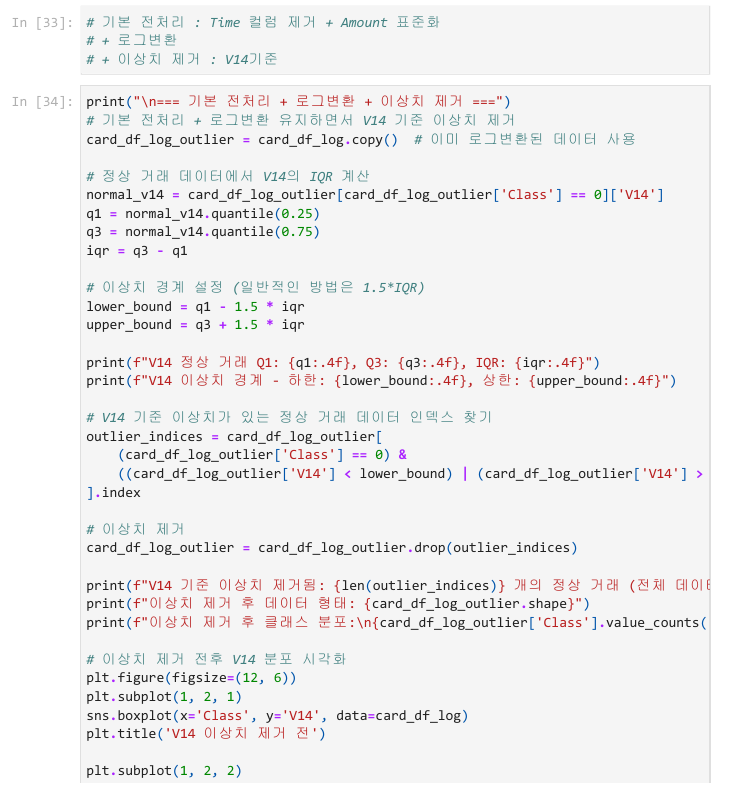
* 표준화로 스케일 통일
* 금액 분포의 왜도를 줄이기 위한 처리
* 로그 변환을 통해 극단적인 고액 거래의 영향력을 완화하고 전체 분포를 정규성에 가깝게 조정

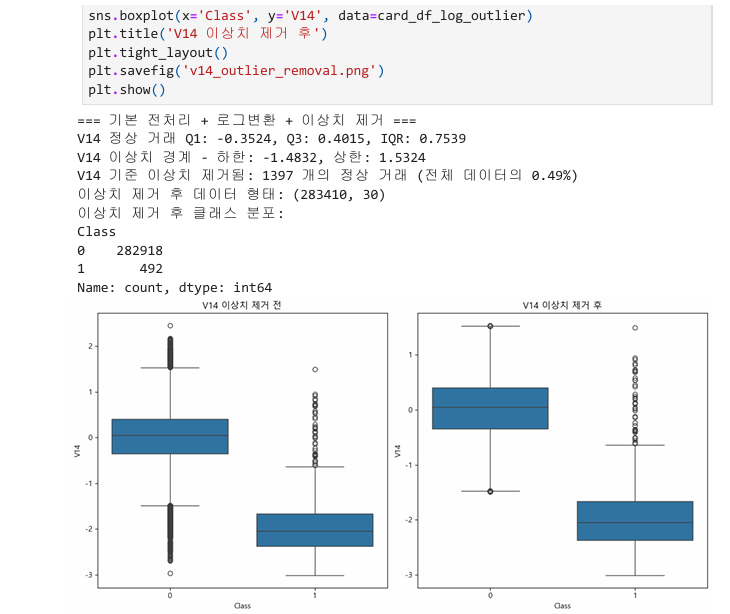




# 이상치 제거 추가 : 기본 전처리 + 로그 변환 + V14 이상치 제거

* V14 피처에서 사기 거래와 분포가 겹치는 일부 정상 거래가 존재
* V14의 IQR 기준 이상치 판단을 통해 약 1100건의 정상 거래를 제거





1. 데이터 분할 및 불균형 처리

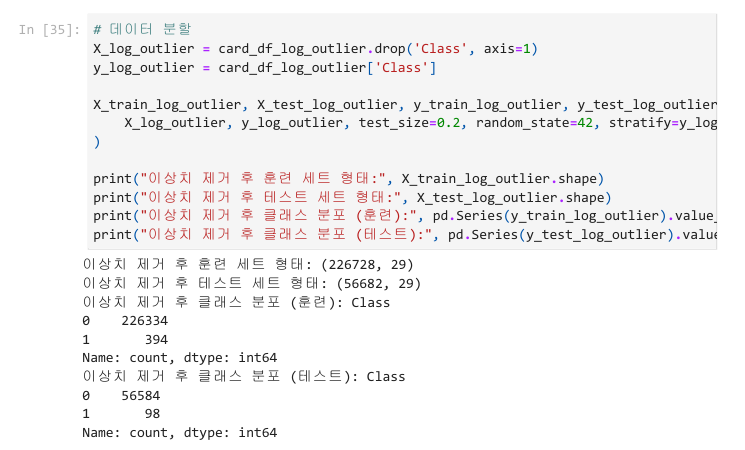
# 4.1 데이터 분할

모델의 성능을 공정하게 평가하기 위해 전체데이터를 훈련셋과 테스트셋으로 분리하였습니다.

* 분할 비율 : 80% 훈련, 20% 테스트
* 랜덤 시드 고정
* 계층 샘플링 적용

→ 사기 거래 비율이 극단적으로 낮기 때문에 훈련셋과 테스트셋 모두에서 클래스 비율이 유지되도록 분할

* 테스트셋은 학습이나 튜닝에 사용하지 않고 최종 성능 평가용으로만 사용했습니다.



# 4.2 불균형 처리

전체 거래 중 사기 거래는 약 0.17%에 불과한 극단적인 클래스 불균형 상황으로

모델이 단순히 ‘정상’만 예측해도 Accuracy가 99% 이상이 되는 문제가 있습니다.

이를 해결 하기 위해 다음 두 가지 전략을 적용했습니다.

1. 클래스 가중치 부여

* 모든 모델에 class\_weight=’balanced’ 설정해 소수 클래스(사기 거래)에 더 높은 가중치를 부여하여 모델이 정상 거래에만 치우치지 않도록 학습을 유도
* 특히 로지스틱 회귀(Logistic Regression)와 LightGBM 모두 클래스 불균형을 내부적으로 보정하는 기능이 있어 클래스 가중치 설정만으로도 일정 수준의 불균형 완화 효과

1. SMOTE

* 훈련셋에만 적용
* 소수 클래스(사기 거래)를 기반으로 인공샘플 생성

→ 사기 거래와 유사한 가상의 거래를 생성하여 샘플 수를 증가시킴

* 결과적으로 훈련셋의 클래스 비율을 1:1에 가깝게 조정
* 테스트셋에는 SMOTE를 적용하지 않음. 현실의 실제 거래 분포를 그대로 반영하기 위해 테스트셋은 원본 상태 유지

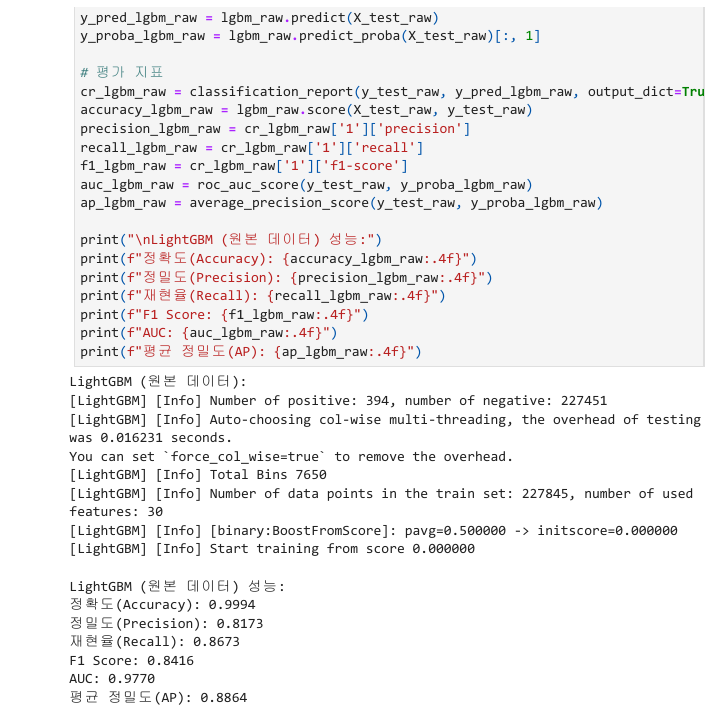
# 4.3 훈련 및 평가 전략

* 각 전처리 조합에 대해 SMOTE 적용 여부(미적용/적용)에 따라  
  총 8개의 훈련 데이터셋을 구성하였고 각 데이터셋에 대해 로지스틱 회귀(Logistic Regression)와 LightGBM 모델을 각각 학습
* 최종적으로 총 16가지 실험 조합을 구성하여 모델 성능 비교 및 전처리 효과 분석을 진행

원본데이터 학습 및 평가

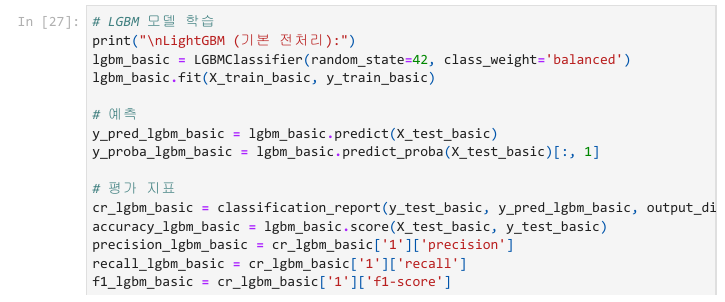


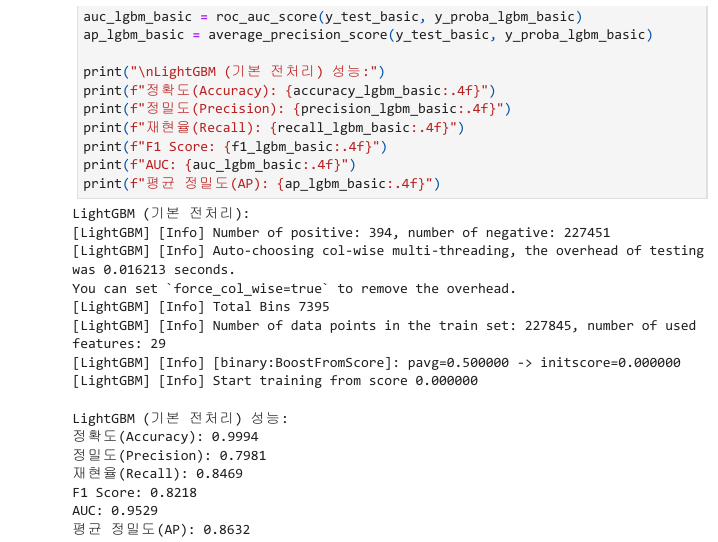




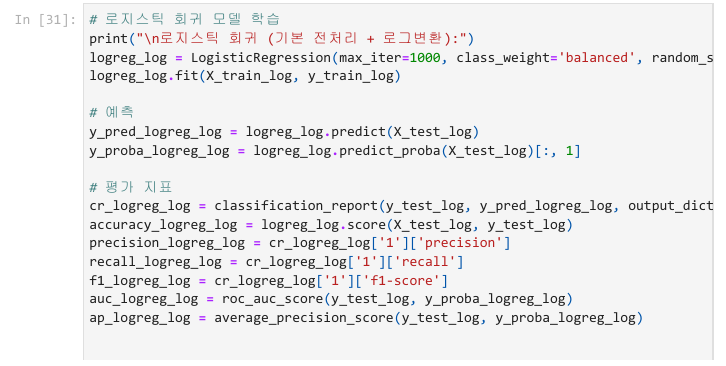
기본 전처리 학습 및 평가

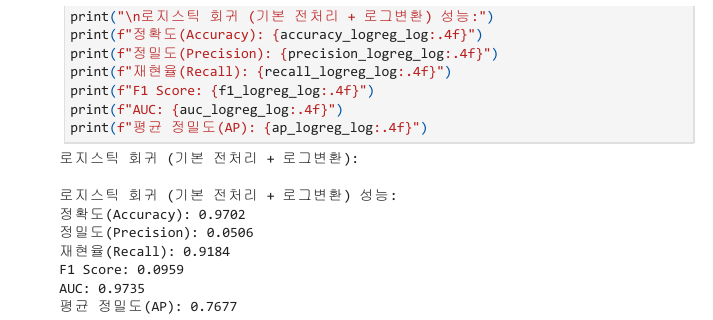




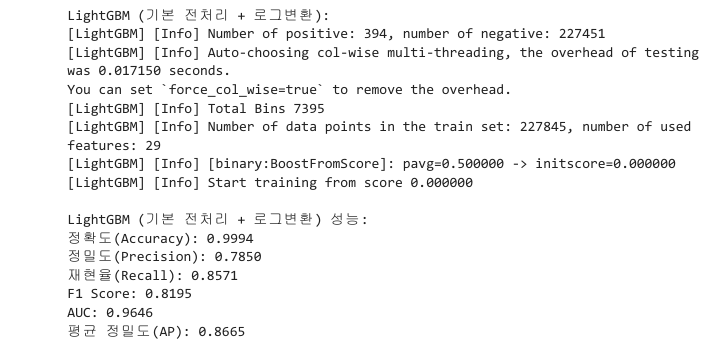


기본 전처리 + 로그 변환 학습 및 평가







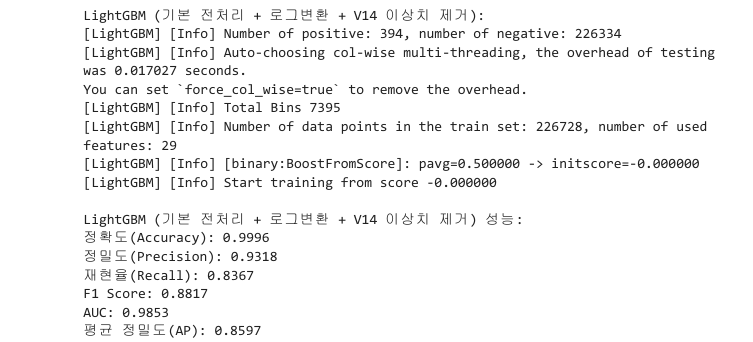


기본 전처리 + 로그변환 + 이상치 제거 학습 및 평가



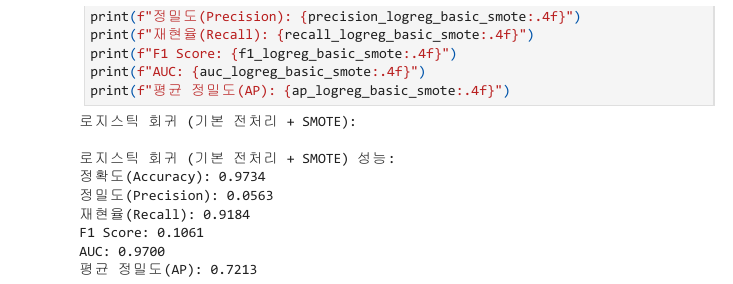


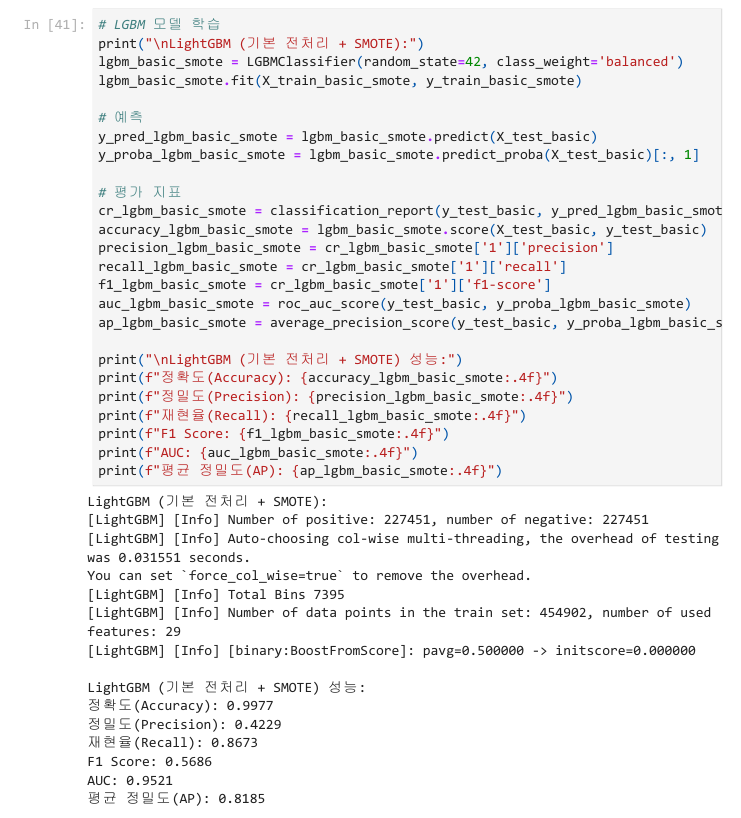




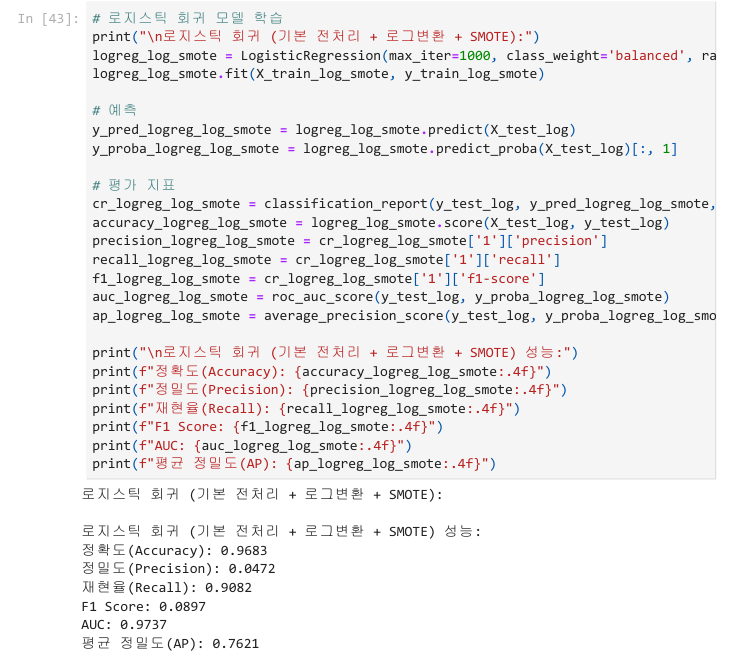
기본 전처리 + SMOTE 학습 및 평가

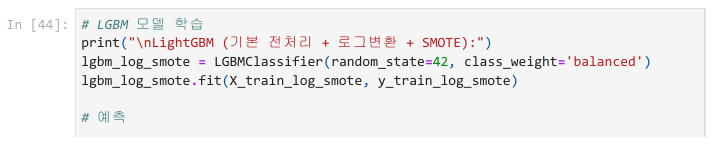


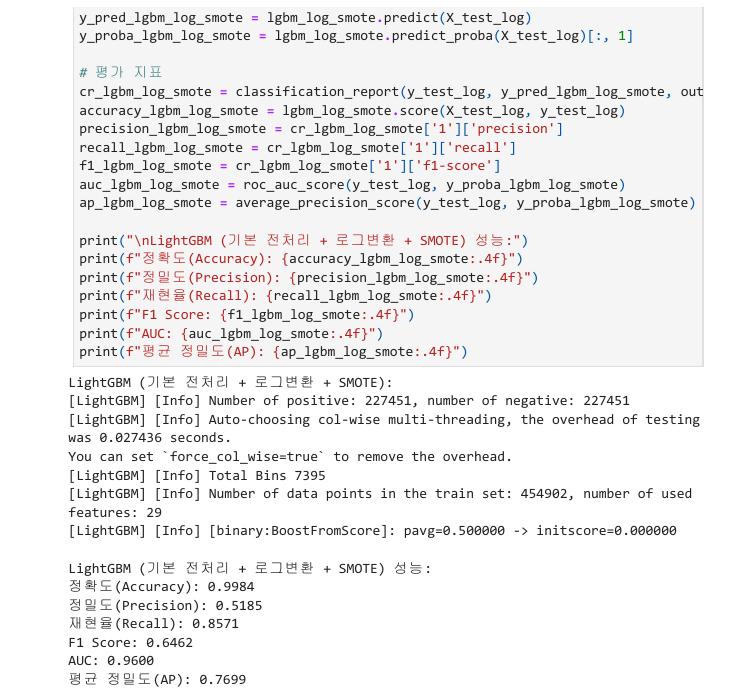




기본 전처리 + 로그변환 + SMOTE 학습 및 평가

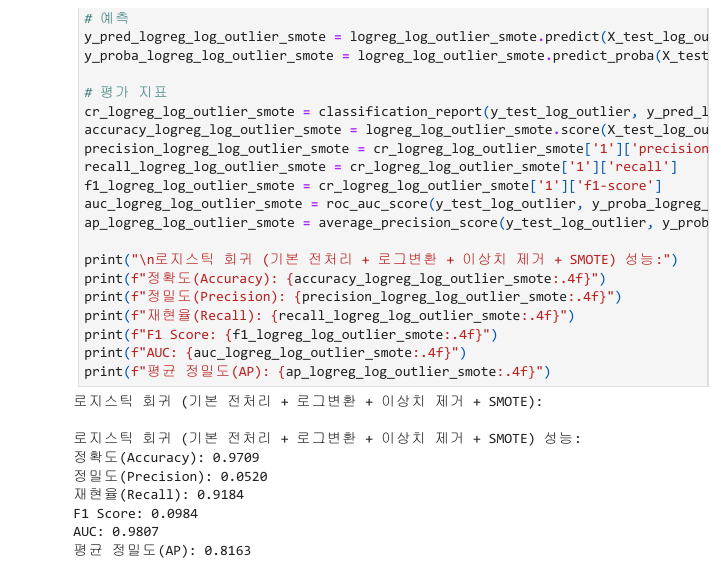






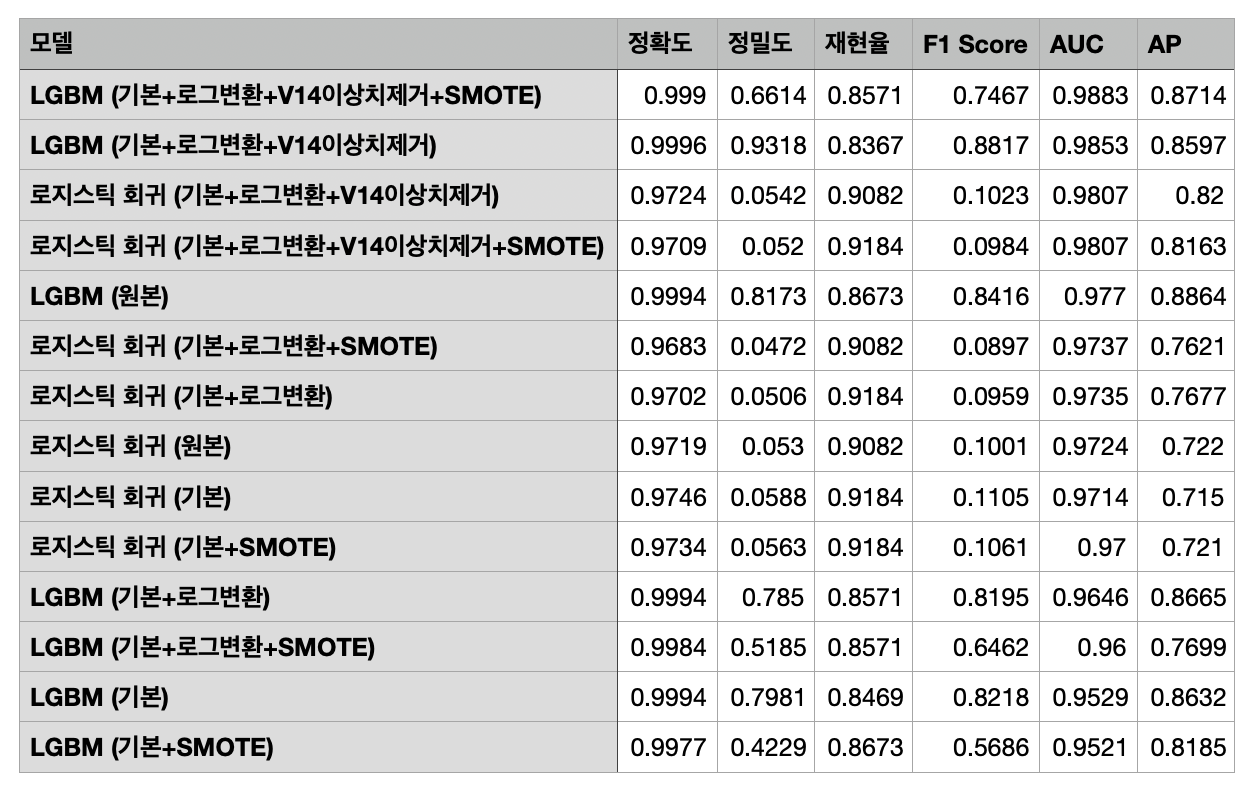
기본 전처리 + 로그변환 + 이상치 제거 + SMOTE 학습 및 평가







1. 모델 학습 및 평가 : 전처리 조합별 비교

각 전처리/샘플링 조합에 대해 로지스틱 회귀와 LightGBM 모델을 학습하고 앞서 분리한 테스트셋을 통해 성능을 평가하였습니다.

표에서 확인할 수 있듯이 전처리 조합에 따라 모델 성능에 유의미한 차이가 나타났습니다.  
특히 LightGBM은 모든 조합에서 로지스틱 회귀보다 F1 Score와 Precision이 월등히 높았으며 기본+로그변환+V14 이상치 제거 조합에서는 F1 Score가 0.88에 달할 정도로 뛰어난 성능을 보였습니다.

반면, 로지스틱 회귀는 모든 조합에서 Precision이 0.05 내외로 매우 낮고 Recall만 과도하게 높은 불균형된 결과를 보여, 실무 적용에는 한계가 있는 것으로 나타났습니다.

또한, SMOTE를 적용한 조합은 전반적으로 Precision 하락을 유발하며 오히려 F1 Score 관점에서는 SMOTE 미적용 조합이 더 우수한 성능을 보였습니다.

이러한 결과를 바탕으로 본 프로젝트에서는 LightGBM + (기본 전처리 + 로그변환 + 이상치 제거) 조합을 최종 모델 후보로 선정하였습니다.

1. 전처리 조합 및 성능 비교 요약

# 6.1 모델 전반 성능

LightGBM은 모든 전처리 조합에서 로지스틱 회귀보다 우수한 Precision과 F1 Score를 보였습니다.

* 로지스틱 회귀는 Precision 5% 내외, F1은 대부분 0.1 미만으로 낮음
* LightGBM은 Precision 40~78%, F1은 최고 0.82까지 도달
* Recall은 두 모델 모두 85~92% 수준이었으나 로지스틱은 Recall을 높이는 대신 False Positive가 많아 Precision 저하로 이어졌습니다

# 6.2 Accuracy 한계

* Accuracy는 두 모델 모두 96~99.9%로 높았지만 불균형 데이터에서 신뢰할 수 없는 지표
* LightGBM은 정상 거래 대부분을 정확히 분류하고, 사기 거래도 상당수 탐지해 높은 Accuracy를 기록
* 모델 비교 시에는 Precision과 Recall 중심으로 해석해야 함

# 6.3 전처리 효과 요약

* 로그 변환은 특히 로지스틱 회귀의 AUC, AP 향상에 효과적 (예: AP 0.72 → 0.77)
* V14 이상치 제거는 LightGBM에서 Precision 향상에 도움
* 예: LightGBM + SMOTE 조합에서 Precision 0.42 → 0.66으로 개선

요약 :   
- 로그 변환 → 선형 모델에 유익  
- 이상치 제거 → 트리 기반 모델 + SMOTE 조합에서 효과적

# 6.4 SMOTE 요약

* 로지스틱 회귀: SMOTE 적용 후 Precision 감소, Recall은 동일 (예: 5.88% → 4.35%)
* LightGBM: SMOTE 적용 시 Recall은 소폭 향상되나 Precision은 크게 하락 (예: 75.0% → 42.3%)
* 결론: 이 데이터에서는 클래스 가중치만으로도 충분했으며 SMOTE는 Precision을 저해하는 경향

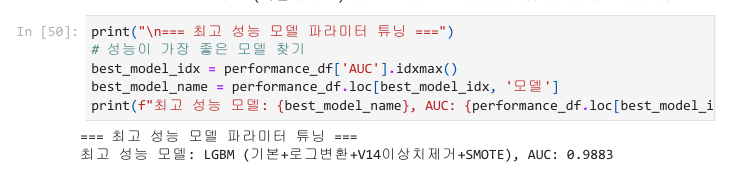
# 6.5 LightGBM vs 로지스틱 회귀

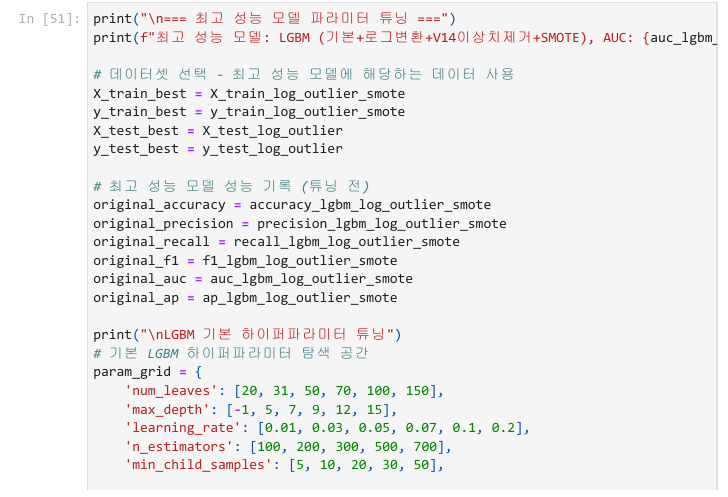
* 로지스틱 회귀는 Recall은 높지만, Precision이 매우 낮아 실무 적용에 부적합
* LightGBM은 적은 오탐지로 높은 검출율을 유지하며, 전체적으로 정밀하고 균형 잡힌 성능을 보임
* 예: Precision 78.5%, Recall 85.7% 조합 → F1 약 0.82

1. 최적 모델 하이퍼파라미터 튜닝

# 7.1 튜닝 조건 및 전략

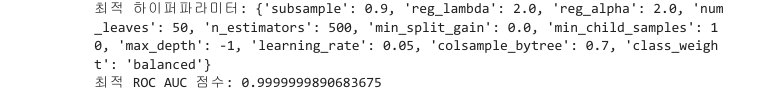
* 전처리 : Time 제거, Amount 로그 변환 + 표준화, V14 이상치 제거
* SMOTE는 미적용
* RandomizedSearchCV로 주요 파라미터(n\_estimators, learning\_rate, num\_leaves 등) 튜닝
* 목표 지표 : ROC AUC



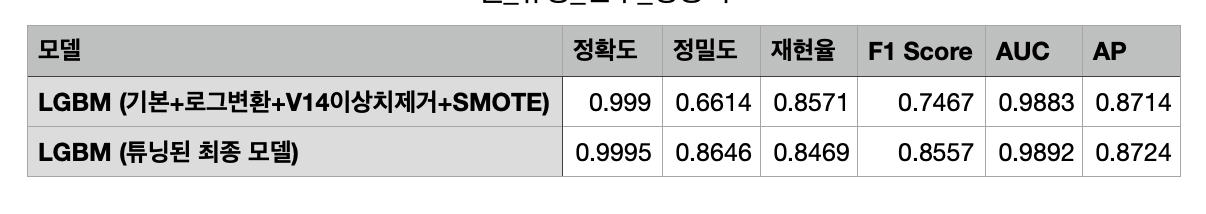




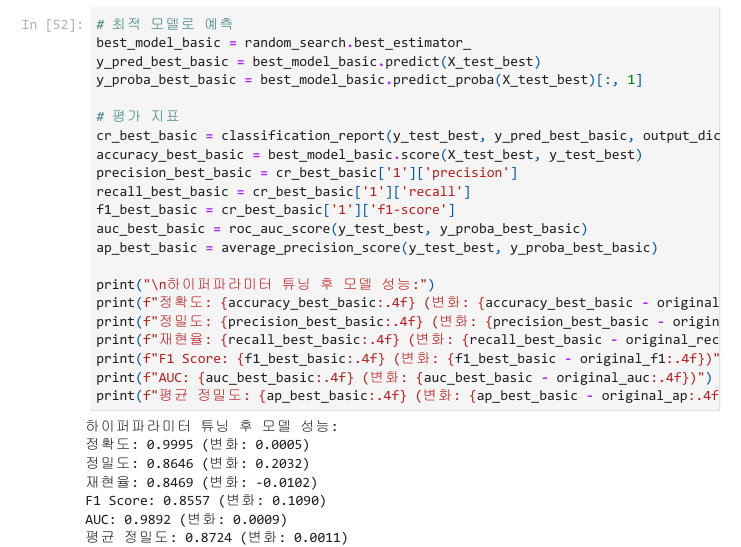




# 7.2 튜닝 결과



* Precision이 크게 향상되어 오탐을 줄이고 Precision-Recall 균형이 개선됨
* Recall은 미세하게 감소했지만 F1 Score는 큰 폭으로 상승 (0.75 → 0.86)



1. 최종 모델 선정 및 적용 논의

# 8.1 최종 모델

본 프로젝트의 최종 모델로는 LightGBM 분류기를 선정하였습니다.  
 적용된 전처리 조합은 다음과 같습니다.

* Time 피처 제거
* Amount 피처에 로그 변환 및 표준화 적용
* V14 피처 기준 이상치 제거
* SMOTE는 미적용 (Precision 저하 방지를 위해)  
  이 조합은 Precision과 Recall 간 균형이 뛰어나고 실제 테스트셋 기준 다음과 같은 우수한 성능을 보였습니다.

# 8.2 모델 적용 가능성

최종 선정된 LightGBM 모델은 실제 신용카드 사기 탐지 시스템에 적용할 수 있을 만큼 실용적이고 우수한 성능을 보입니다.

Precision은 86.5%, Recall은 84.7%로 대부분의 사기 거래를 포착하면서도 정상 거래에 대한 오탐지를 최소화합니다. 이는 오탐지(정상 거래 차단)와 미탐지(사기 거래 누락) 간의 균형이 잘 잡힌 모델임을 의미합니다.

또한, AUC 0.9892의 높은 분류 성능을 기반으로 분류 임계값(threshold)을 조정하여 Precision과 Recall 간의 비율을 상황에 맞게 유연하게 조절할 수 있습니다. 예를 들어, 오탐을 줄이고자 할 경우 threshold를 높이고, 사기를 놓치지 않으려면 낮춰서 Recall을 우선시할 수 있습니다.

LightGBM은 예측 속도가 빠르고 경량화된 모델로, 실시간 거래 탐지 시스템에도 적용 가능합니다. 사용된 PCA 기반 특성은 실제 환경에 맞게 원본 피처나 새로운 특성 엔지니어링으로 대체할 수 있습니다.

LightGBM은 SHAP, LIME과 같은 해석 도구와도 호환성이 높아 예측 결과에 대한 설명 가능성도 확보할 수 있습니다.

따라서 본 모델은 정확도, 실시간성, 해석 가능성, 유지보수 측면 모두에서 뛰어나며 실제 금융 현장에서도 충분히 활용 가능한 수준의 모델로 판단됩니다.

1. 결론 및 인사이트

이번 프로젝트는 극도로 불균형한 금융 데이터를 기반으로, 실제 금융사에서 중요한 과제인 신용카드 사기 거래 탐지 문제를 머신러닝으로 해결하고자 한 실전형 프로젝트입니다. 단순한 정확도(Accuracy) 이상의 정교한 지표(F1 Score, Precision, Recall, AUC 등)를 기반으로 다양한 모델과 전처리 전략을 조합하여 실험을 수행하고 실제 운영 환경에 적용 가능한 최적의 모델을 도출하는 것이 주요 목표였습니다.

# 9.1 결론 요약

* 모델 성능 측면
* LightGBM 모델이 모든 조합에서 가장 안정적이고 우수한 성능을 보였습니다.
* 특히 전처리 조합(로그 변환 + 이상치 제거)과 함께 사용한 튜닝된 LightGBM은  
   F1 Score 0.85 이상, Precision 86.5%, Recall 84.7%, AUC 0.989로 매우 높은 성능을 달성했습니다.
* 반면, 로지스틱 회귀는 Recall은 높지만 Precision이 매우 낮아 실무 적용에는 부적절했습니다.
* 전처리 전략 효과
* Amount 로그 변환과 V14 기준 이상치 제거는 전체적으로 모델 성능을 개선하는 데 효과적이었습니다.
* 특히 이상치 제거는 Precision 향상에 크게 기여했습니다.
* 불균형 처리 전략
* 단순 클래스 가중치 조정만으로도 충분한 성능 확보가 가능했고 SMOTE는 오히려 Precision을 떨어뜨리는 부작용을 야기하여 실제 운영에서는 신중히 고려해야 할 기법임을 확인했습니다.
* 튜닝 성과
* 하이퍼파라미터 튜닝을 통해 F1 Score가 약 +0.10 이상 향상, Precision이 +20% 이상 상승하였습니다.
* 모델의 균형 잡힌 예측 능력을 더욱 극대화할 수 있었으며 사기 거래에 대한 탐지 정확도와 신뢰도가 높아졌습니다.

# 9.2 실무 인사이트

모델의 성능은 단순히 정확도(Accuracy)만으로 평가할 수 없습니다. 특히 본 프로젝트처럼 클래스 불균형이 심한 경우에는 Precision, Recall, AUC, PR-AUC 등의 다양한 지표를 함께 고려해야 실제로 유용한 모델을 만들 수 있습니다. 이때 중요한 것은 오탐지(False Positive)와 미탐지(False Negative)의 균형 조절입니다. 사기 거래를 놓치지 않기 위해서는 높은 Recall이 필요하지만, 동시에 경고의 오작동을 줄이기 위해서는 Precision도 확보되어야 하므로 이 두 지표는 본질적으로 상충하는 관계에 있습니다. 따라서 분류 임계값(threshold)을 조정하거나, 현업의 정책적 목표에 맞춘 튜닝이 반드시 병행되어야 합니다.

또한, 모델의 해석 가능성과 운영 안정성도 실무 적용에서 매우 중요한 요소입니다. LightGBM 모델은 높은 예측 성능뿐만 아니라 피처 중요도 분석이 가능하여, 모델이 어떤 요인을 근거로 판단했는지를 파악할 수 있습니다.  
더불어, 예측 속도가 매우 빠르기 때문에 실시간 거래 탐지 시스템에도 적용하기 적합합니다.

이 외에도, 전처리 전략은 모델 성능 향상에 있어 결정적인 역할을 합니다. 예를 들어, Amount에 로그 변환을 적용하거나 V14 기준으로 이상치를 제거하는 등의 간단한 처리만으로도 Precision이 20~30% 가까이 향상되는 결과를 확인할 수 있었습니다.  
이러한 전처리의 영향은 결과 해석에서 반드시 고려되어야 할 핵심 요소입니다.