5-1 ML 과제 - 신용카드 사기 탐지 머신러닝 모델링: 클래스 불균형 해결 및 성능 최적화

작성자: 28기 남건우

1. 개요

본 분석의 목적은 신용카드 거래 데이터셋을 활용하여 사기 거래(Fraud)를 탐지하는 분류 모델을 구축하는 것입니다. 특히, 사기 거래 데이터가 매우 적은 '클래스 불균형' 문제를 해결하기 위해 SMOTE 기법을 적용하고, 비즈니스 관점에서 중요한 지표인 Recall(재현율)과 F1-score, PR-AUC를 극대화하는 것을 목표로 합니다.

2. 데이터 전처리 및 EDA

데이터 구성 - creditcard.csv 데이터를 활용하였으며, 총 31개의 피처와 Class 라벨로 구성됨을 확인할 수 있었습니다.
클래스 분포 - 정상 거래(0) 대비 사기 거래(1)의 비율이 극히 낮음을 확인하여 불균형 처리의 필요성을 확인하였습니다.
변수 변환 - 다른 변수에 비해 단위 차이가 큰 Amount 변수에 대해 StandardScaler를 적용하여 표준화를 수행하였습니다.
데이터 분할 - 학습셋과 테스트셋을 8:2 비율로 분할하되, stratify 옵션을 사용하여 클래스 비율을 일정하게 유지하였습니다.

3. 방법론

3.1 클래스 불균형 해소: SMOTE

적용 이유: 사기 거래 클래스 비율이 샘플링 이전 0.9983:0.0017, 샘플링 이후 0.9531: 0.0469로 사기 거래 데이터 비율이 매우 낮습니다. 사기 거래 데이터가 부족한 상태에서 학습할 경우, 모델이 다수 클래스인 정상 거래에만 편향되어 실제 사기 거래를 탐지하지 못하는 문제가 발생합니다. SMOTE는 소수 클래스 주변에 가상의 데이터를 생성하여 모델이 사기 거래의 패턴을 충분히 학습할 수 있도록 돕습니다. SMOTE를 통해 학습 데이터 내 사기 거래 건수를 정상 거래 건수와 비슷한 수준으로 증폭시켜 모델의 변별력을 확보할 수 있습니다.

3.2 모델 선정

Random Forest는 다수의 결정 트리를 결합하는 앙상블 기법으로, 오버피팅에 강하고 클래스 불균형 데이터에서도 안정적인 성능을 보여 선정하였습니다.

4. 분석 결과

모델 학습 후 테스트 데이터셋(2,099건)에 대해 평가를 수행한 결과는 다음과 같습니다.

4.1 지표별 결과

- Class 0 - 사기 거래 기준

Metric	Result	Target	Status
Recall (재현율)	1.00	>= 0.80	달성
F1-score	1.00	>= 0.88	달성
PR-AUC	0.95	>= 0.90	달성

- Class 1 - 사기 거래 기준

Metric	Result	Target	Status
Recall (재현율)	0.89	>= 0.80	달성
F1-score	0.92	>= 0.88	달성
PR-AUC	0.95	>= 0.90	달성

4.2 상세 평가 해석

Precision(정밀도): 정상 클래스 0.99 / 사기 클래스 0.95로, 사기라고 예측한 경우 실제 사기일 확률이 매우 높았습니다.

Recall (재현율): 정상 클래스 1.00 / 사기 클래스 0.89로, 실제 사기 거래의 약 89%를 성공적으로 탐지함.

PR-AUC: 0.95라는 높은 수치를 기록하며, 불균형 데이터셋에서도 정밀도와 재현율이 조화롭게 작동함을 증명하였습니다.

5. 최종 결론 및 제언

목표 달성 여부: 분석 결과 요구되는 모든 지표에서 과제 목표치를 달성하는 좋은 성능을 거두었습니다. 특히 사기 탐지의 핵심인 Recall이 0.80을 크게 넘어서며 실전 투입이 가능한 수준의 탐지 성능을 확보하였습니다.

성능 향상 요인으로는 1. SMOTE 적용을 통해 소수 클래스에 대한 학습량을 증대한 점과 2. Random Forest의 앙상블 효과로 인해 데이터의 노이즈에 강건하게 반응한 점을 뽑을 수 있습니다.

추후 제언으로 몇가지를 말씀드리자면 우선 단일 테스트셋 평가를 넘어 Stratified K-Fold 교차 검증(Cross-Validation)을 수행함으로써, 데이터의 특정 구간에 과적합되지 않은 보편적인 성능의 일관성을 검증할 필요가 있습니다. 또한, 실제 금융 현장에서는 '사기를 놓치지 않는 재현율'과 '정상 고객을 오인하지 않는 정밀도' 사이의 기회비용이 상충하므로, Precision-Recall Curve 분석을 바탕으로 비즈니스 목적에 최적화된 Threshold를 세밀하게 조정하는 프로세스가 병행되면 좋을 것 같습니다.