

# 펀딩피 이상거래 탐지 함수

1. 각 피쳐의 상관관계를 알기

1. 시트 가공 : 계정 단위로 df 만들기

- 펀딩피 절댓값

## 1) funding\_fee\_abs\_mean

$$f_1 = \frac{\text{mean}(|\text{funding\_fee}|) - \min}{\max - \min}$$

필요데이터: Funding ['funding\_fee']

이미지 삭제 왜 안되지 → ㅋㅋㅋㅋㅋㅋㅋㅋㅋ

- 포지션 보유시간

## 2) mean\_holding\_minutes

(holding은 작을수록 이상 → 역정규화)

$$f_2 = 1 - \frac{\text{mean\_holding}}{\max(\text{holding})}$$

account\_id + position\_id

필요데이터: Trade ['openclose', 'ts']

- 펀딩 시작 거래 집중도

## 3) funding\_timing\_ratio

$$f_3 = \text{ratio\_of\_trades\_within\_funding\_time}$$

본래 0~1 사이이므로 정규화 불필요.

필요데이터: Funding['ts']

- 펀딩피 수익 비중

## 4) funding\_profit\_ratio

$$f_4 = \frac{\text{funding\_profit}}{\text{total\_profit}}$$

필요데이터: account\_id 별 Funding['funding\_fee'],

Trade['amount']

최종 df만들기

account_id	f1_funding_fee_abs_mean	f2_mean_holding_minutes	f3_funding_timing_ratio	f4_funding_profit_ratio	f1
A_A312...	2.11	12.5	0.54	0.82	0
B_91AD...	0.5	215	0.12	0.02	0
...					

상관관계 분석



#### ✓ A. mean\_holding\_minutes ↔ funding\_profit\_ratio = **0.529**

👉 보유 시간이 짧을수록 펀딩피 수익 비중이 늘어난다

(= 전형적인 펀딩 헌터 패턴)

➡ 가중치 가장 크게 줘야 함 (0.35 ~ 0.4 추천)

#### ✓ B. funding\_fee\_abs\_mean은 surprisingly 낮은 상관관계

- 원인은 펀딩 데이터 자체에 정상적 거래도 large funding\_fee가 발생할 수 있기 때문
- 절댓값 자체는 noise가 커서 단독 feature로는 weak signal

➡ 중간 가중치 (0.15~0.2 정도)

#### ✓ C. funding\_timing\_ratio도 상관관계 낮음

- 유저별 거래 횟수 차이가 크기 때문
- 하지만 도메인 지식 상 “중요한 feature”  
(상관은 낮아도 feature로 포함해야 한다)

➡ 중간~낮은 가중치 (0.1~0.15)

#### ✓ D. funding\_profit\_ratio는 피처 중 가장 직관적으로 의미가 큼

상관값만 보면 -0.529(holding)과 높지만

이 값은 “핵심 결과”를 나타내므로 매우 중요.

➡ 가장 큰 가중치 (0.35~0.4)

## 🔥 최종 가중치 추천 (Funding Hunter 전용)

Feature	Weight	이유
funding_profit_ratio	<b>0.40</b>	펀딩 수익에 대한 의존 → 가장 직접적 신호
mean_holding_minutes	<b>0.30</b>	짧을수록 펀딩 시각 스나이핑 패턴 강화
funding_timing_ratio	<b>0.15</b>	펀딩 시각 거래 집중 → 보조 신호
funding_fee_abs_mean	<b>0.15</b>	펀딩피 자체 규모 → noise가 있으나 의미 있음

최종 합수 식:

$$\text{FundingHunterScore} = 0.40 \cdot FPR + 0.30 \cdot (1 - HM) + 0.15 \cdot FTR + 0.15 \cdot FFA$$

여기서:

- $FPR$  = funding\_profit\_ratio\_norm
- $HM$  = mean\_holding\_minutes\_norm
  - 보유시간이 짧을수록 위험이므로 “ $1 - HM$ ” 사용
- $FTR$  = funding\_timing\_ratio\_norm
- $FFA$  = funding\_fee\_abs\_mean\_norm