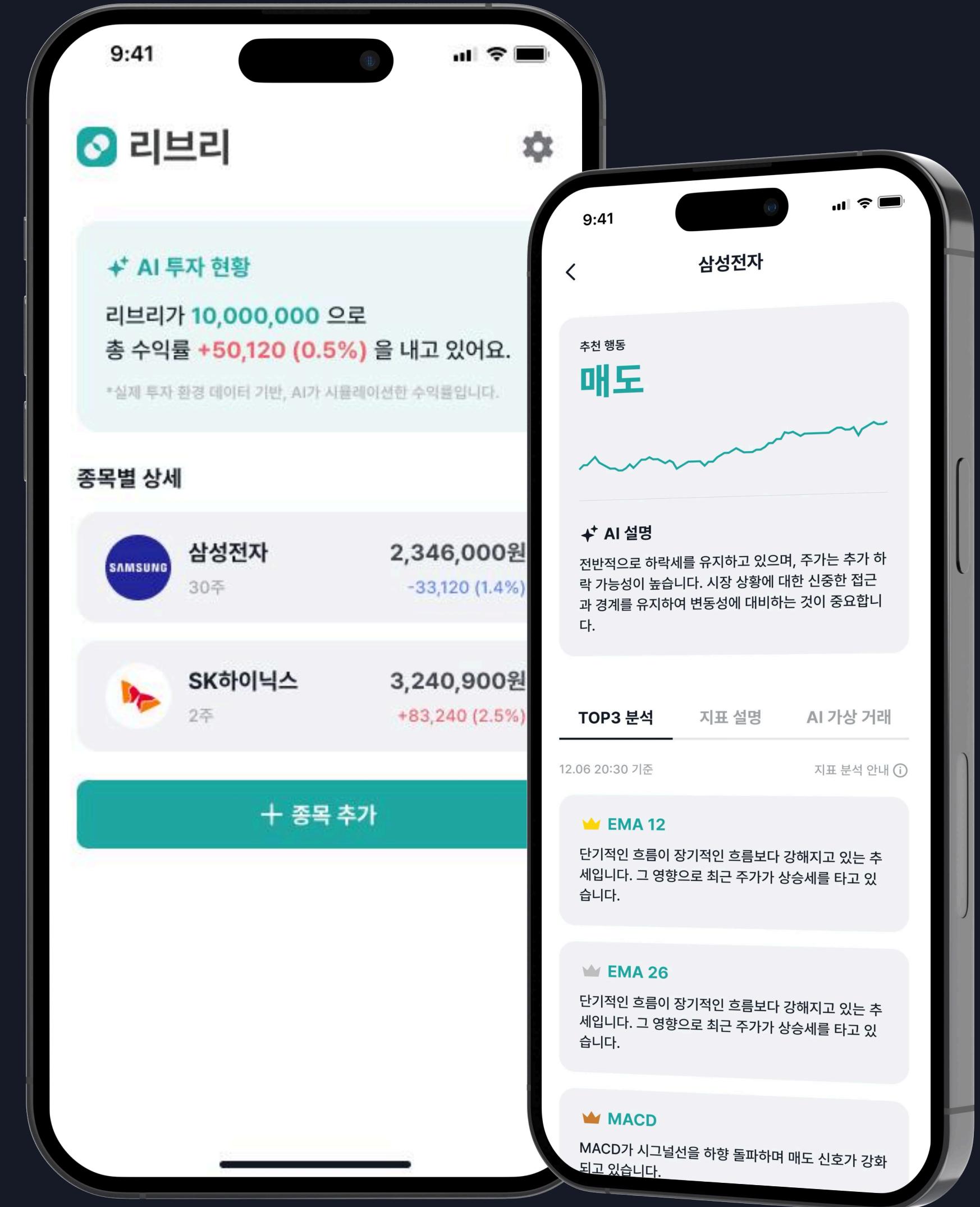


강화학습 기반 AI 투자 보조 플랫폼



팀 리브리



주식 투자

경험이 있으신가요?

성인 인구의 35%

국내 주식 시장 개인투자자

1,410만 명

거래회전율

연 185.4%

(미국 110%, 일본 108.9%)

신규 유입 투자자 **60% 이상**이 손실을 경험

개인투자자는 **시장수익률 하회**하는 투자성과 기록

개인투자자들은 여러 *행태적 편의로 인해, 비효율적인 투자행태를 보이고 있다.

*행태적 편의 : 투자의사결정 과정에서 경험, 감정, 신념, 선호 등의 영향으로 비합리적인 투자행태가 유발되는 심리적 기제

과잉확신

손실회피

처분효과

복권형 주식 선호

친숙성 편향

단기군집거래

투자 판단의 기준이 일관되지 않고, 검증되지 않은 정보와 감정에 의해 흔들린다.

금융위원회 산하 공식 연구기관인 **자본시장연구원(KCMI)**에 따르면,

1

전문가 판단

개인의 직관에 의존한
의사결정으로 편향이 강화



전문가 판단을 참고할
객관적 기준 제공 필요

2

행동 변화 중심 교육 콘텐츠

단순 지식 교육만으로는
편향이 교정되지 않음



'왜 이렇게 판단했는지'
이해하는 행동 기반 교육 필요

3

정보 형식/구조/투명성 개선

판단 근거가 보이지 않으면
단편적·감정적 매매가 발생



의사결정 과정이 드러나는
투명한 정보 제공 필요

1

차트, 지표 제공형 서비스



2

AI 매매, 예측 서비스



3

자문 서비스



결과가 아닌 과정을 학습하는 투자 보조 플랫폼, 리브리

왜 이런 판단이 나왔는지



외부 지표 + TOP3 분석

👑 EMA 12

단기적인 흐름이 장기적인 흐름보다 강해지고 있는 추세입니다. 그 영향으로 최근 주가가 상승세를 타고 있습니다.

👑 EMA 26

단기적인 흐름이 장기적인 흐름보다 강해지고 있는 추

판단 기준을 학습하고



지표 설명

EMA12

최근 12일간의 주가 평균을 계산하되, 최근 가격에 더 큰 비중을 두는 이동평균선입니다. 주가의 단기적인 흐름을 빠르게 파악하는데 사용됩니다.

💡 해석 포인트

- 주가가 EMA12 위에 있으면 단기 상승 흐름으로 볼 수 있습니다.

결과로 복기한다.



AI 가상 거래

오늘

10주 판매

+12,830원 (22.3%)

1주당 98,750원

어제

리브리에 사용된 강화학습 모델

A2C

Advantage Actor-Critic

MARL

Multi-Agent RL

단일 에이전트 기반의
빠른 반응형 모델

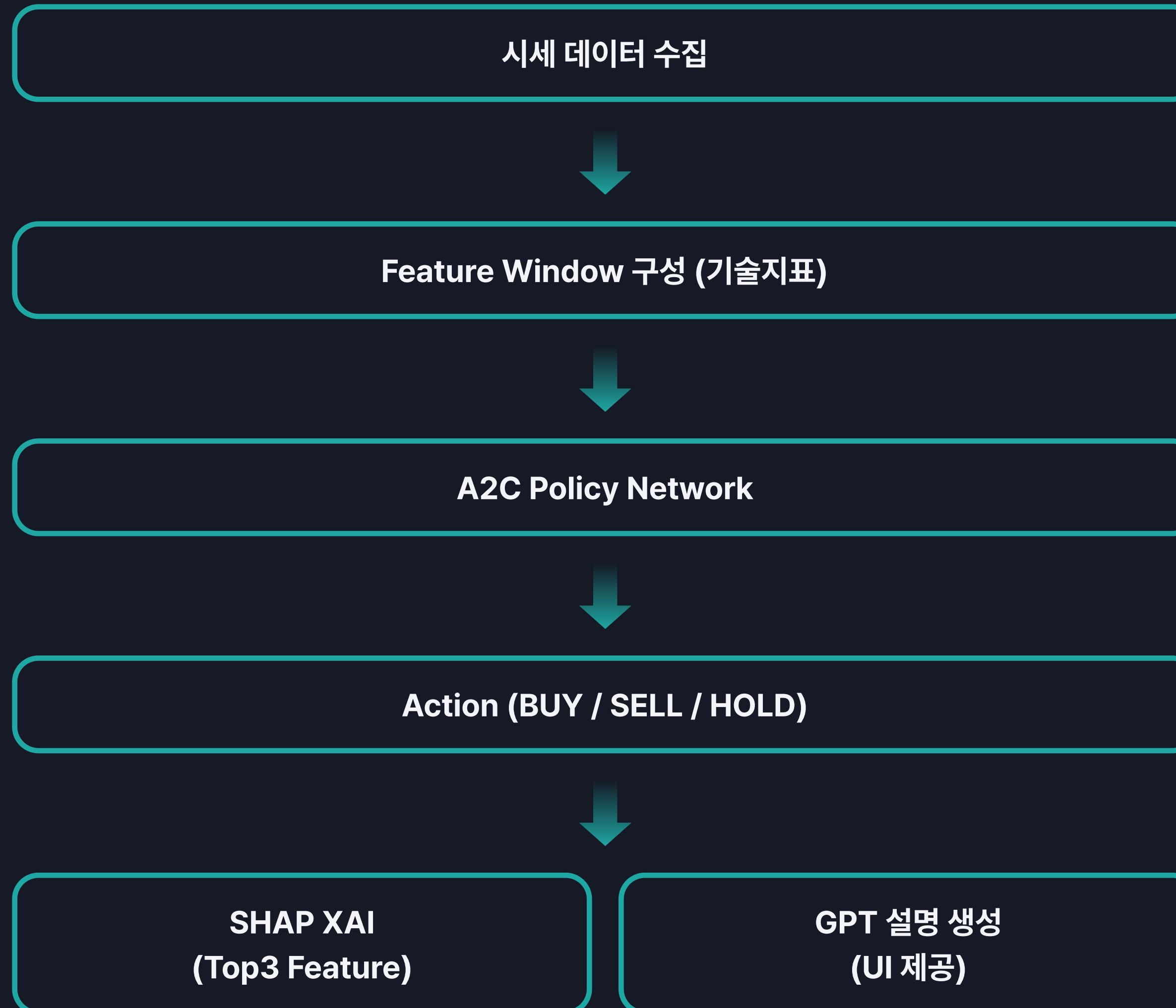
다중 에이전트 협력 기반의
안정적 의사결정 모델

A2C 모델 개요

- ▷ Actor(정책) + Critic(가치평가) 병렬 학습 구조
- ▷ 단일 에이전트 기반 **단기 시그널 탐지**에 강점
- ▷ 변동성 높은 시장에서 매수/매도 의사결정 품질 확보
- ▷ 리브리 서비스 내에서 **공격 성향 포트폴리오** 담당



“빠르게 시장 방향을 포착하고
즉시 의사결정 가능한 모델”



Key Features

- ▷ Window 기반 State 입력으로 시계열 특성 반영
- ▷ Temperature Sampling으로 확률 기반 행동 결정
- ▷ SHAP Top-3 지표 추출로 해석 가능성 (Explainability) 보장

추론 처리 흐름

▷ `load_model()`

모델 및 스케일러 리소스 로딩

▷ `predict_today()`

`last_window` 생성 후 `run_a2c_decision()` 호출

▷ `_select_action()`

Softmax 확률 기반 행동 샘플링

▷ **Result Generation**

Top-3 지표에 Raw Value 매핑 + GPT 설명

“A2C가 어떻게 행동을 고르고,
왜 그 행동을 했는지 한번에 확인”

A2C_core_logic.py

Python 3.11

```
def run_a2c_decision(last_window):
    # 1. 상태 정의 (State Construction)
    state = build_state(last_window, position_flag=0)

    # 2. A2C 정책 네트워크 추론 → 확률 기반 행동 결정
    with torch.no_grad():
        s_t = torch.tensor(state, dtype=torch.float32).unsqueeze(0)
        logits, _ = agent.ac_net(s_t)
        action, probs = _select_action_from_logits(
            logits,
            temperature=0.8,
            min_conf=0.45,
            min_margin=0.10,
            sample=True,
        )

    # 3. XAI - Top-3 중요 지표 추출
    _, _, top_features = explain_a2c.get_top_features(
        state, agent, explainer, feature_names, top_k=3
    )

    # 4. Raw 값 + 중요도를 GPT 프롬프트를 딕셔너리로 정리
    feature_importance, technical_indicators = {}, {}
    for i, feat in enumerate(top_features):
        name = (feat.get("name") or feat.get("base")
                or feat.get("indicator") or f"feature_{i}")
        if "importance" in feat:
            feature_importance[name] = float(feat["importance"])
        if "value" in feat:
            technical_indicators[name] = float(feat["value"])

    # 5. GPT로 지언어 설명 생성
    explanation = interpret_model_output(
        signal=ACTION_ID_TO_EN[action],
        technical_indicators=technical_indicators,
        feature_importance=feature_importance,
    )

    return action, probs, top_features, explanation
```

A2C 최종 수익률
+100.3%

액티브 전략 (Active Strategy)

단순 보유 수익률
+83.9%

패시브 벤치마크 (Buy & Hold)

코스피 지수 수익률
+63.5%

시장 지수 (Market Index)

A2C 백테스트 성과 비교

● A2C 모델 ● 단순 보유 ● KOSPI

포트폴리오 가치 (원)

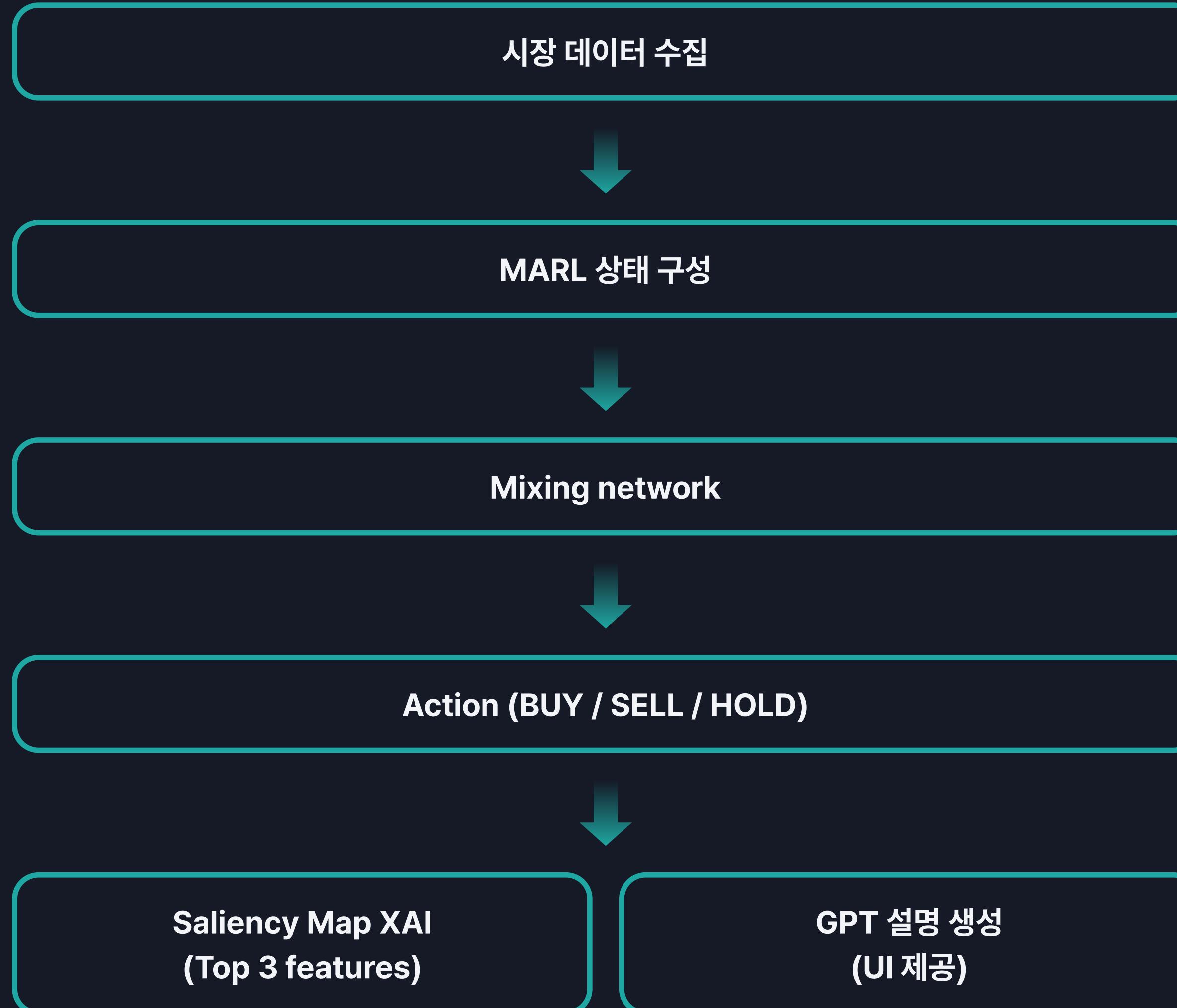


MARL 모델 개요

- ▷ 3개의 독립적인 AI 트레이더(Agent)가 팀을 이뤄 투자 결정
- ▷ 학습은 중앙에서 통합적으로, 실행은 각자 독립적으로 수행하는 CTDE 방식
- ▷ 개별 행동이 팀 전체 수익에 기여하는 정도를 정교하게 산출하는 QMIX 알고리즘
- ▷ 리브리 서비스 내에서 안정 성향 포트폴리오 담당



“3개의 에이전트 의견을 다수결로 결정해 안정성을 높인 모델”



Key Features

- ▶ 투표를 통해 최종 의사결정을 내리므로, 한쪽으로 치우치지 않는 균형 잡힌 투자
- ▶ Saliency Map기반 XAI로 해석 가능성 (Explainability) 보장

추론 처리 흐름

▷ load_model()

DataProcessor 로딩 및 QMIX Learner 초기화

▷ predict_today()

오늘 기준 데이터 재가공 및 MARL 환경 설정

▷ Agent Loop

각 에이전트별 행동(Softmax) 및 중요도 산출

▷ Joint Action & XAI

다수결 투표로 Signal 생성 + Top-3 근거 추출

“다중 에이전트(QMIX)가 협력하여
최적의 공동 행동(Joint Action)을 도출”

A2C_core_logic.py

Python 3.11

```
def run_a2c_decision(last_window):
    # 1. 상태 정의 (State Construction)
    state = build_state(last_window, position_flag=0)

    # 2. A2C 정책 네트워크 추론 + 확률 기반 행동 결정
    with torch.no_grad():
        s_t = torch.tensor(state, dtype=torch.float32).unsqueeze(0)
        logits, _ = agent.ac_net(s_t)
        action, probs = _select_action_from_logits(
            logits,
            temperature=0.8,
            min_conf=0.45,
            min_margin=0.10,
            sample=True,
        )

    # 3. XAI - Top-3 중요 지표 추출
    _, _, top_features = explain_a2c.get_top_features(
        state, agent, explainer, feature_names, top_k=3
    )

    # 4. Raw 값 + 중요도를 GPT 프롬프트용 딕셔너리로 정리
    feature_importance, technical_indicators = {}, {}
    for i, feat in enumerate(top_features):
        name = (feat.get("name") or feat.get("base")
                or feat.get("indicator") or f"feature_{i}")
        if "importance" in feat:
            feature_importance[name] = float(feat["importance"])
        if "value" in feat:
            technical_indicators[name] = float(feat["value"])

    # 5. GPT로 자연어 설명 생성
    explanation = interpret_model_output(
        signal=ACTION_ID_TO_EN[action],
        technical_indicators=technical_indicators,
        feature_importance=feature_importance,
    )

    return action, probs, top_features, explanation
```

MARL 최종 수익률

+100.5%

QMIX Agent (Active)

단순 보유 수익률

+83.9%

패시브 벤치마크 (Buy & Hold)

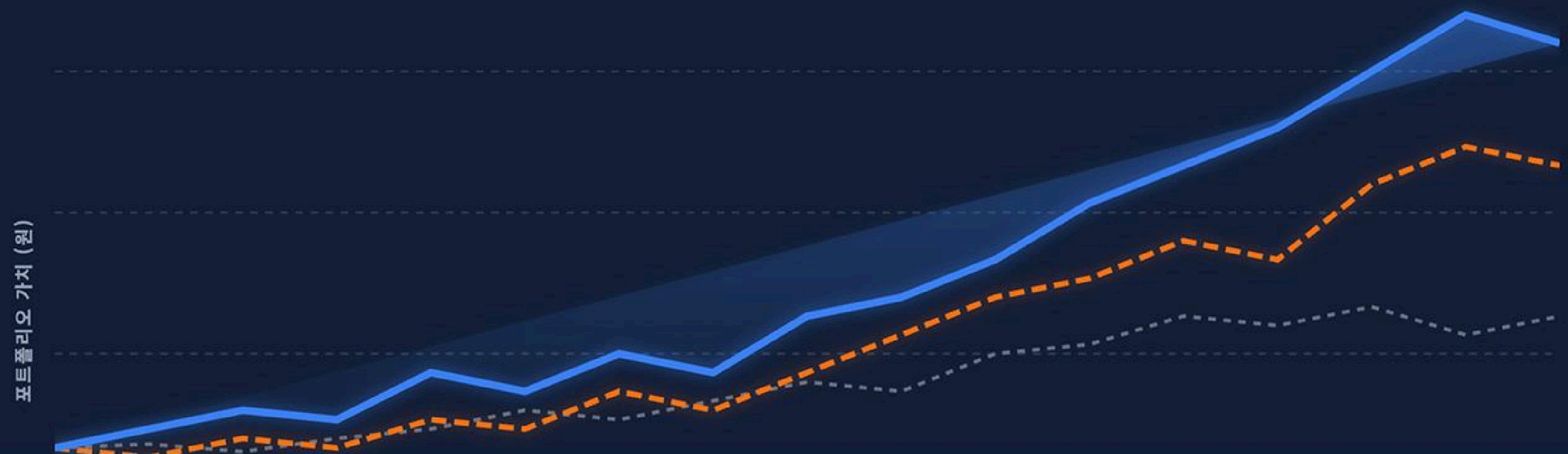
코스피 지수 수익률

+63.5%

시장 지수 (Market Index)

MARL (QMIX) 백테스트 성과 비교● QMIX Agent ● 단순 보유 ● KOSPI

포트폴리오 가치 (원)



기대효과

리브리는 투자 판단 과정을 학습시키는 구조 자체를 제안하는 프로젝트입니다.

개인 투자자 학습 효과

판단 근거를 이해하는
투자 학습 경험 제공

XAI 기반 지표 설명으로 금융
리터러시 및 의사결정 역량 강화

AI 가상 거래를 통한 시행착오
비용 감소 및 감정적 매매 완화

산업적 기여 및 확장성

주식 외 ETF, 가상자산 등 다양한
금융 상품으로 확장 가능

외부 API가 아닌, 직접 설계 및 학습한
강화학습 모델과 XAI 구조를 기반으로 한
금융 AI 시스템으로 확장 가능

오픈소스 및 학술적 기여

강화학습 + XAI 결합 사례로서
연구 및 교육용 레퍼런스 제공

AI 의사결정 설명 구조에 대한 실증적
프로젝트 사례 축적



결과가 아닌 과정을 학습하는 투자 플랫폼