**SK네트웍스 Family AI과정 12기  
 모델링 및 평가 수집된 데이터 및 전처리 문서**

**□ 개요**

* 산출물 단계 : 모델링 및 평가
* 평가 산출물 : 수집된 데이터 및 전처리 문서
* 제출 일자 :
* 깃허브 경로 : https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN12-FINAL-2TEAM
* 작성 팀원 : 윤 권

## 1. 데이터 수집 및 구성

#### 본 프로젝트는 Yahoo Finance API를 활용하여 거래량 기준 상위 100개 종목의 주가 데이터를 수집하였다. 수집 범위는 최근 2년간의 일봉(Daily) 데이터이며, 각 종목별로 다음과 같은 정보를 포함한다: Open, High, Low, Close, Volume, Adj Close.

#### 이러한 데이터는 주식 시장의 가격 움직임과 거래 강도 등 기술적 분석에 활용되는 주요 지표로서, 향후 딥러닝 및 강화학습 모델의 입력 값으로 사용된다.

## 2. 데이터 전처리

### 2.1 전처리 개요

#### 효과적인 시계열 예측 및 강화학습 학습을 위해, 수집된 주가 데이터를 여러 단계의 전처리를 거쳐 모델 입력에 적합한 형태로 변환하였다.

### 2.2 결측치 처리 (주말, 휴일 등 휴장일)

#### 주가 데이터는 주말, 공휴일, 천재지변, 거래정지 등으로 인해 일부 날짜에 결측이 발생한다. 이러한 날짜는 모든 종목에서 동일하게 비어 있기 때문에, 전체 종목에서 해당 날짜 자체를 제거함으로써 시계열 정렬을 유지하였다.

#### 또한, 특정 종목이 상장된 지 얼마 되지 않아 존재하지 않는 기간이 있을 경우, 해당 종목의 앞부분을 잘라내어 전체 종목과 시간축을 맞추었다.

##### 이와 같은 방법은 시계열의 연속성과 정합성을 유지하기 위한 핵심 과정이다.

### 2.3 이상치 처리

#### 비정상적으로 높은 변동을 보이는 일봉 데이터는 (High – Low) / Close > 0.3 기준으로 이상치로 간주하였다. 이는 스플릿, 급등락, 미체결 급등락 주문 등 외부 요인에 의한 비정상 데이터일 가능성이 높기 때문이다.

#### 처리 방식은 아래와 같다.

##### 윈저라이징(Winsorizing): 상/하위 1% 구간으로 잘라 Clipping

##### 필요시 해당 날짜 제거

### 2.4 노이즈 제거 (Noise Reduction)

#### 시계열 데이터는 미세한 단기 변동성(노이즈)가 많아, 학습 안정성에 부정적인 영향을 줄 수 있다.

#### 이를 완화하기 위해 아래와 같은 방법을 적용하였다.

##### 이동평균(SMA) 필터: Close 가격에 3일 또는 5일 이동평균 적용

##### 기술적 지표 기반 변동성 판단(Bollinger Band)으로 이상구간 보정

#### 이러한 기법은 LSTM의 학습 효율을 높이고, 강화학습 에이전트가 의미 있는 패턴에 집중할 수 있게 도와준다.

### 2.5 정규화 및 스케일링

#### 정규화는 입력값의 범위를 일정하게 맞춰주어, 모델이 특정 feature에 편향되지 않도록 하는 데 필수적이다.

#### 본 프로젝트에서는 다음과 같은 스케일링 기법을 혼합하여 적용하였다.

#### 로그 수익률 변환

##### 가격 정보(Open, High, Low, Close)는 비정상성을 제거하기 위해 로그 수익률로 변환

#### Min-Max 정규화

##### 기술적 지표 등 일부 비정규화 데이터는 [0, 1] 범위로 변환

#### StandardScaler (Z-score표준화)

##### 거래량(Volume)과 같이 편차가 큰 변수는 평균 0, 표준편차 1 기준으로 표준화

#### 정규화는 시계열 특징에 따라 피처별로 적절한 방식으로 개별 적용되었으며, 전체 시계열 구조를 유지하기 위한 후처리도 병행되었다.

### 2.6 슬라이딩 윈도우 & 상태 구성

#### 본 프로젝트에서 강화학습 환경은 슬라이딩 윈도우 방식으로 구성된다.

#### 각 State는 최근 5~7일치 시계열 데이터를 포함하며, 이 데이터를 기반으로 LSTM 모델이 상태 벡터를 생성한다.

#### 하나의 에피소드는 4개의 연속된 State로 구성되며, 각 State마다 에이전트는 매수(Buy), 매도(Sell), 관망(Hold) 중 하나의 행동을 선택한다.

#### 이러한 구성은 단기적 시장 추세를 기반으로 전략을 수집하는 데 적합하며, 강화학습의 학습 속도 및 안정성에 긍정적인 영향을 준다.

### 3.7 학습/검증 데이터 분할

#### 시계열 데이터의 시간 의존성을 고려하여, 랜덤 분할이 아닌 시간 기준 분할을 적용하였다.

#### 학습 구간은 2023년 7월부터 2025년 1월까지, 검증 구간은 2025년 2월부터 2025년 7월 까지로 설정하였다.

#### 이러한 분할 방식은 미래 데이터를 예측하는 실제 투자 상황과 유사한 형태로 평가를 수행할 수 있게 해준다.

## 4. 모델 구조 및 학습 방식

### 4.1 LSTM Encoder

#### LSTM(Long Short-Term Memory)은 시계열의 장기 의존 관계를 잘 학습할 수 있는 RNN 계열 모델로, 본 프로젝트에서는 각 슬라이딩 윈도우(5~7일치 데이터)를 입력으로 받아 이를 요약한 상태 벡터(State Vector)를 출력한다.

#### 이 상태 벡터는 강화학습 에이전트의 입력으로 활용되어, 과거 시장 흐름을 반영한 의사 결정을 가능하게 만든다.

### 4.2 DQN Agent

#### 강화학습 에이전트는 DQN(Deep Q-Network)을 기반으로 하며, LSTM이 출력한 상태 벡터를 입력으로 받아 현재 시점에서 최적의 행동을 선택한다.

#### 행동 공간은 매수, 매도, 관망 3가지로 구성되어 있으며, 에피소드 종료 시점에서의 누적 수익률을 보상(Reward)으로 받아 학습이 이루어진다.

## 5. 보상(Reward) 설계 및 수수료 반영

### 5.1 보상 설계 기준

#### 강화학습 에이전트의 보상은 에피소드 종료 후 누적 수익률을 기준으로 계산된다. 보상의 단위는 수익률 1% 1점으로 설정되었으며, 이는 다음과 같이 계산된다.

### 5.2 수수료 반영

#### 실제 거래를 반영하기 위해, 매 거래 시 0.5%의 수수료가 발생하도록 설정하였다. 즉, 포지션 진입 시 0.5%, 청산 시 0.5%로 총 1.0%의 거래 비용이 보상 계산에 반영된다.

#### 예시

##### 수익률 +3.0%, 거래 발생 → 수수료 1.0% 차감 → 최종 수익률 +2.0% → Reward = 2.0

##### 수익률 -1.0%, 거래 발생 → 수수료 1.0% 차감 → 최종 수익률 -2.0% → Reward = -2.0

#### 이로 인해 에이전트는 수익을 극대화함과 동시에 불필요한 거래를 최소화하는 전략을 학습하게 된다.

## 6. 향후 개선 방향

#### 실적 발표, 금리 인상, CPI 등 이벤트 기반 외부 변수를 피처로 통합하여 현실성과 대응력 강화

#### DQN 외에 PPO(Proximal Policy Optimization), SAC(Soft Actor-Critic) 등 다양한 강화학습 알고리즘과 비교 실험 수행

#### 단순 수익률 중심 보상 외에도, 샤프지수(위험 조정 수익률), 최대 낙폭(MDD) 등을 고려한 복합 보상 함수 설계