Project Live-Large

210220

목표:

1. 주어진 시계열 데이터를 통해 학습하여 ~~다음 시각의 가격~~다음 시점까지의 가격 변동률을 예측하는 모델 구축
2. Auto-regressive모델로 발전시켜 장기적 가격 형성 시뮬레이션
3. 모델 정확도 검증 후 HTS와 연동시켜 자동 매매

모델 개요

* 적용 범위: 각 종목 전용 모델
* 아키텍쳐 선정: 딥러닝 기반
  + seq2one
    - Linear/CNN
  + seq2seq
    - RNN
    - Transformer
  + Transformers
* 학습 방식
  + 준비: 배치 학습
  + 운용: 온라인 학습
    - lr설정
* 모델 입력
  + 주의점
    - 시간에 대한 로버스트 성: 단기적 변동성과 장기적 동향을 동시에 파악할 수 있어야 한다.
      * 단기성은 S에, 장기성은 E에 적용
    - 다양한 경제 요소의 반영
      * E에 반영
  + 크기: (입력 시계열 길이S, 배치 크기 N, 특징량 수 E)
    - 입력 시계열 길이 S
      * 단위: 일/주(/월)
      * 예: 연속 30일, 연속 30주
    - 배치 크기 N: 초기 학습 데이터로 과거 얼마 정도의 기간을 사용할 것인가?
      * 예: 1년, 3년
    - 특징량 수 E: 종목에 따라 유연하게 설정
      * Raw data
        + 주식 시장의 매력도

국내 시장: KOSPI/KOSDAQ

해외 시장:

미국 국채 10년물 금리

금값

* + - * + 섹터의 매력도

국내 섹터 지수: 업종별 등락(NAVER 참조하여 직접 설계 혹은 섹터 추종 지수)

관련 원자재 가격

* + - * + 종목의 특성

주가(시/저/고/종)

거래량

투자자

개인 보유비율

기관 보유비율

외국인 보유비율

* + - * 전처리
        + 기본 지수: 산술(%)이 아닌 기하(로그)적 접근

종목

주가 변화율: log(당일 종가/전일 종가)

거래량 변화율: log(당일 거래량/전일 거래량)

장내 격동성: log(당일 고가/당일 저가) → 하지만 개장 중 주로 머물렀던 가격대는?

장외 격동성: log(당일 시가/전일 종가)

섹터

섹터 지수 변화율: log(당일 마감치/전일 마감치)

섹터 대비 종목의 변동성: log(주가 변화율 / 섹터 지수 변화율)

시장

국내 시장 대비 종목의 변동성: 주가 변화율 / KOSPI 변화율

금리 변화율: log(폐장 시 금리/개장 시 금리)

금리 격동성: log(전일 폐장 시 금리/당일 폐장 시 금리)

* + - * + 이동평균

5: 최단기

20: 단기

60: 중기

120: 장기

240: 최장기

* + - * + 변화율
* 모델 출력
  + 크기: (출력 시계열 길이T, 배치 크기 N, 특징량 수 E)