

[Day 01] 2020-01-06

개별 주식의 모형 vs 전체를 총괄하는 모형

- 개별 모형을 주는 것이 성능은 더 좋다.
- 서브프라임 모기지론.... 10년치 데이터가 더 안좋다. -> 오히려 Outlier... -> 유의미한 결론 5년이내로!

Facebook prophet

- 페이스북에서 개발한 시퀀스 모델 알고리즘
- 체크포인트를 나누어 각 구간별 선형 회귀 적합한 모델
- 장점 : 빠른 학습시간, 쉬운 사용법, 특정 이벤트 기간(슈퍼볼, 선거일 등등) 설정 가능
- 단점 : 찾아낸 바로는 Input이 Date와 Close만 가능...?, 파라미터 조정이 어려움.

새로운 지표를 만드는 것이 가장 중요!!!

- RSI, MACD, 볼린저 밴드...!
- 전일의 open, high, low, close를 Input으로 Linear Regression을 적용한 뒤, 그 Output을 LSTM의 Input으로 적용
- 상관관계가 높은 Top10과 낮은 Bottom10의 종목을 가중치를 부여하고, Input으로 사용

2000개의 종목을 하나의 모형으로

- 10%는 거래 정지 종목...
- 현실 반영

새로운 모델 방향

- prophet의 결과를 Input으로, 상한선과 하한선도 Input으로...!
- 각 회사들의 상관관계를 반영!!!

최종방향

- 지금 상황에서 의미가 불분명하고 불확실한 Input을 추가하기 보다는 모델을 개선하는 방향으로 나아가기로함.
- 현재 LSTM 단일 모델에서 ARIMA, LSTM, GRU, PROPHET 등의 4가지 모형을 사용해서 각 모형을 비교 및 각모형의 결과값들을 평균을 한다는지 새로운 Stacking 모델을 구축.
- 목표는 5일 뒤 주가의 정확한 가격 예측
- 각자 역할
- 박상희 : LSTM
- 김민기 : PROPHET
- 우수민 : GRU
- 박장순 : ARIMA