



High Frequency Trading survey

Tae-Won Lee,
leetee95@gmail.com
2021. 08. 31.

Introduction

- High Frequency Trading(HFT)
 - HFT(고빈도 트레이딩)는 특수한 컴퓨팅 장비를 이용하여 분,초 단위의 거래로 수익을 내는 트레이딩 방식
 - 한국의 경우 미국(0.1%) 등에 비해 높은 증권 거래세(0.25%)로 인해 HFT가 불가능한 국가로 여겨짐
 - 2023년 증권 거래세 개편(0.25% -> 0.15%)으로 HFT가 가능해질 것으로 예상.

Introduction

- High Frequency Trading(HFT)
 - Trading에서 향후 주가 예측은 중요한 문제가 될 수 있음
- ex) 예측한 주가의 하락이 예측된다면 매도 주문, 예측한 주가의 상승이 예측 된다면
매수 주문을 하는 전략 가능.[1]

Introduction

- Issue

- HFT에서 모델에서 예측의 중요성을 말하는 논문은 많았지만, 효율성을 높이기 위한 논문 탐색의 어려움

Related works

- High frequency trading with deep learning

학회 명	IF	논문 명	발간 년도	인용 수	내용
ACM Transactions on Management Information Systems	2.312	A Double-Layer Neural Network Framework for High-Frequency Forecasting	2017	38	1) Dnn 을 High-Frequency data에 적용함 2) High frequency data에서 종목간의 상관성을 발견, 특정 주식을 예측하는데 다른 주식 데이터 이용 3) 주가 예측에 도움이 되는 기술적 지표 이용
Journal of Forecasting	1.57	An ensemble of LSTM neural networks for high-frequency stock market classification	2018	64	1) 기술적 지표들을 사용한 lstm ensemble model 이용하여 분류
Quantative finance	1.491	Encoding of high-frequency order information and prediction of short-term stock price by deep Learning	2019	25	1) 호가창의 정보를 이용하기 위해 CNN을 이용하여 feature 탐색하여 neural network의 input으로 이용 2) 다음 주가의 상승/하락을 분류함
Journal of Computational Science	2.644	Sequence classification of the limit order book using recurrent neural networks	2018	63	1) 호가창데이터를 이용해 sequence의 등락을 분류
Mathematical Problems in Engineering	1.009	Stock Market Prediction on High-Frequency Data Using Generative Adversarial Nets	2018	98	1) Generative Adversarial Nets을 이용하여 1분봉의 종가 예측

Related works

- High frequency trading with deep learning

학회 명	IF	논문 명	발간 년도	인용 수	내용
journal of risk and financial management	0.44	Trend Prediction Classification for High Frequency Bitcoin Time Series with Deep Learning	2019	31	1) 분 단위 데이터의 가상화폐의 상승과 하락을 분류하는 모델
journal of risk and financial management	0.44	Deep LSTM with Reinforcement Learning Layer for Financial Trend Prediction in FX High Frequency Trading Systems	2019	21	1) Lstm과 강화학습을 이용한 hft 시스템 구축 2) 환율데이터(fx)를 이용

- 호가창 데이터를 이용한 상승 하락 분류
- 기술적 지표들을 이용한 예측력 향상

Related works

- Dataset

글로벌 선물(지수 상품, 금, 천연가스), 외환(Fx)거래, 암호화폐 등이 트레이딩에 좋은 종목으로 여겨짐.

이유

- 1) 정치적인 규제 요인에 의한 영향이 적음
- 2) 단일 주식에 비해 유동성이 크고 변동성이 적음 (암호화폐 제외)
- 3) 신뢰가 가능한 전통적인 매매시장
- 4) 전세계의 모든 투자자들이 참여하는 큰 규모의 시장이므로 특정 세력에 의한 조작이 불가능

>> 해외선물 데이터는 680\$ 의 월 수수료를 지불한 후, 증권사에서 api에서 다운로드 가능 (비용 지불 문제)

>> 환율 분봉 데이터는 제공되지 않음

Related works

- Dataset

암호화폐 데이터

```
2020-12-31 23:23:00 28255.559000 28255.559000 ... 0.003165 89.416129
2020-12-31 23:26:00 28255.559000 28255.559000 ... 0.001214 34.298858
2020-12-31 23:27:00 28255.558320 28255.558320 ... 0.007483 211.435213
2020-12-31 23:32:00 28213.175000 28415.960839 ... 0.006875 195.311767
2020-12-31 23:33:00 28346.965518 28346.965518 ... 0.000589 16.596629

[100000 rows x 6 columns]
```

API를 이용해 데이터를 불러온 화면

- 1) 업비트 api를 이용해 데이터 습득 가능
- 2) 증권사 api가 32비트 환경을 제공, 암호화폐 api는 64비트 제공.
- 3) 한번에 10만개 이상의 분봉 데이터 호출 가능
- 4) 24시간 동안 연속되는 시장
(하루가 지남에 따라 시계열이 분할되는 현상)



장이 끝나고 시작되는 지점(1분봉, 일반 주식)

Related works

- Dataset

학회 명	IF	논문 명	발간 년도	인용 수	사용 데이터	비고
The Journal of Financial Data Science		Deep Reinforcement Learning for Trading	2021	41	Fx, index	
Computational Finance	1.34	Agent Inspired Trading Using Recurrent Reinforcement Learning and LSTM Neural Networks	2017	33	USD/GBD	We take 2000 data points from US Dollar/British Pound price series, we used 1000 30 minutes interval price points from 1/6/17 to 2/3/17 as training set and again take 1000 30 minutes interval price points from 2/3/17 to 3/6/17 as test set.
Acm transactions on information systems	2.889	A Double-Layer Neural Network Framework for High-Frequency Forecasting	2017	39	S&P 500 (January 1, 2013 to May 31, 2013)	5-minute interval, large dataset that includes 100 stocks with the largest capitals in the S&P 500.
Journal of Forecasting	1.57	An ensemble of LSTM neural networks for high-frequency stock market classification	2018	64	22 large cap US stocks	We apply our methodology to 22 large cap US stocks. We use 1 year of raw trade data, which were cleaned and aggregated into 5-minute intervals, amounting to roughly 19,000 observations per stock.
International Conference on Computer Science & Education (ICCSE)	0.47	High-Frequency Stock Trend Forecast Using LSTM Model	2018	14	CSI 300	
journal of risk and financial management	1.557	Trend Prediction Classification for High Frequency Bitcoin Time Series with Deep Learning	2019	21	BTCUSD, BTCCNY	

Related works

- Dataset

학회 명	IF	논문 명	발간 년도	인용 수	사용 데이터	비고
journal of risk and financial management	0.44	Deep LSTM with Reinforcement Learning Layer for Financial Trend Prediction in FX High Frequency Trading Systems	2019	21	EUR/USD, GBP/USD, EUR/GBP	1분 간격
Quantitative Finance	1.491	financial Trading as a Game: A Deep Reinforcement Learning Approach	2018	24	Fx	

Related works

- 호가창 데이터 이용

- 실시간 호가 데이터 이용가능
- 과거 호가 데이터는 실시간 데이터의 수집이 필요

호가 데이터 수집 예시: **증권의 api를 활용하여 2019년 3월 20일부터 2019년 5월 7일까지 코스닥 (KOSDAQ) 시가총액 상위 50종목의 실시간 호가 잔량 정보를 수집

- 방법 1. 2개월간 데이터 수집과 동시에 일반적인 1분봉 데이터로 진행 중 2개월 후 데이터 교체
- 방법 2. 호가 데이터를 수집하지 않고 1분봉 데이터로 진행

- 실시간 데이터 사용 여부

- 트레이딩 구현이 아닌 모델의 성능평가만을 위해서는 실시간 데이터가 필요하지 않을 것으로 예상.

Related works

- High frequency trading을 평가하는 방식

- 1) Classification : 다음 time step의 상승, 하락 등을 분류하여 분류의 정확성 계산 [2], [3].

- 2) back testing을 통한 수익률 계산.

- ex) We can see Buy and Hold achieves the best profit rate followed by LSTM-Individual, Random and Optimistic. That is because LSTM-Individual, Random and Optimistic are high-frequency trading strategies with naturally low profit rate, based on short-term trend while Buy and Hold strategy only trades once. [1]

- 위 논문에서 언급 된 1) Buy and Hold, 2) LSTM-Individual 3) Random, 4) Optimistic 전략 등을 이용하여 back testing을 진행하여 각 전략의 수익률을 판단 가능.

Reference

- 1 Siyu Yao, Linkai Luo and Hong Peng. High-Frequency Stock Trend Forecast Using LSTM Model. [International Conference on Computer Science & Education \(ICCSE\)](https://doi.org/10.1109/ICCSE.2018.8468703). <https://doi.org/10.1109/ICCSE.2018.8468703>
- 2 [Svetlana Borovkova](#), [Ioannis Tsiamas](#). An ensemble of LSTM neural networks for high-frequency stock market classification. Journal of Forecasting (2019). <https://doi.org/10.1002/for.2585>
- 3 Takuya Shintate and Lukáš Pichl. Trend Prediction Classification for High Frequency Bitcoin Time Series with Deep Learning. Journal of Risk and Financial Management (2019). <https://doi.org/10.3390/jrfm12010017>