인공지능을 이용한 비트코인 매매

Bitcoin Trading with Artificial Intelligent

연구자

성복고등학교(10618) 송주영

**목차**

1. 서론
2. 이론적배경
3. 비트코인 데이터 예측 모델 구축 및 실험

3-1알고리즘 구축 및 구현

3-2 데이터수집 및 처리

3-3매매기록 및 설명

1. 평가&결론
2. 실험환경
3. **서론**

연구를 시작할 당시에 코인을 활용해서 돈을 버는 투자가 유행이었다. 그러나 아무 전략없이 하는 투자는 돈을 잃기도 쉬웠고 한강 간다는 말을 커뮤니티에서 자주 볼 수 있게 되었다. 우리는 이러한 투자 실패에 대비해 과학, 수학의 힘을 빌리고자 했다. 따라서 우리는 RNN(Recurrent Neural Network, 순환신경망)의 일종인 LSTM[[1]](#footnote-1)(Long Short Term Memory)를 활용한 자동매매장치를 개발하여 수익을 극대화하는 방법을 선택했다.



(관계없는 사진(한강)

**키워드: LSTM, 비트코인, 블록체인, 선물거래**

1. **이론적 배경**

비트코인이란- 비트코인은 사토시 나카모토 라는 가명을 쓰는 프로그래머에 의해 만들어 져서 2009년에 배포가 되었다. 중앙은행이 없어 사람 대 사람(P2P)으로 거래가 진행되어 시장이 비교적 자유롭고 거래 수수료 또한 적다는 특징이 있다. 해킹을 방지하기 위해 블록체인 [[2]](#footnote-2)기술을 사용해 사용자들의 컴퓨터에 각각 나누어 거래 내역을 저장한다. 현재 엘살바도르라는 국가에서 법정화폐로 쓰고 있다.



출처: http://www.coinreaders.com/669

**비트코인 선물거래란?**

일반적으로 선물은 일정시점에 현재가격으로 살수 있는 약속을 하는 개념이다.

현재 거의 모든 상품에 선물 시장이 존재한다. 하지만 이러한 선물거래는 증거금의 10%만 납부하면 가능하기 때문에 일종의 레버리지 역할을 한다. 바이낸스의 선물거래는 내가 가지고 있는 돈에 레버지리를 일이 켜서 즉 돈을 빌려서 투자한다는 개념이다.

예를 들어 떡상이라는 사람의 현재 돈이 1만원이 있다.

이 떡상이라는 사람이 비트코인이 올라갈 것을 예측하고 비트코인 레버리지 100배 롱포지션을 잡았다 이럴 경우 떡상이가 산 실제 금액은

1만원 \* 100 = 100만원이 되는 것이다.

100만원에서 1%만 떨어져도 1만원이므로 이때 청산이 발생하게 된다.

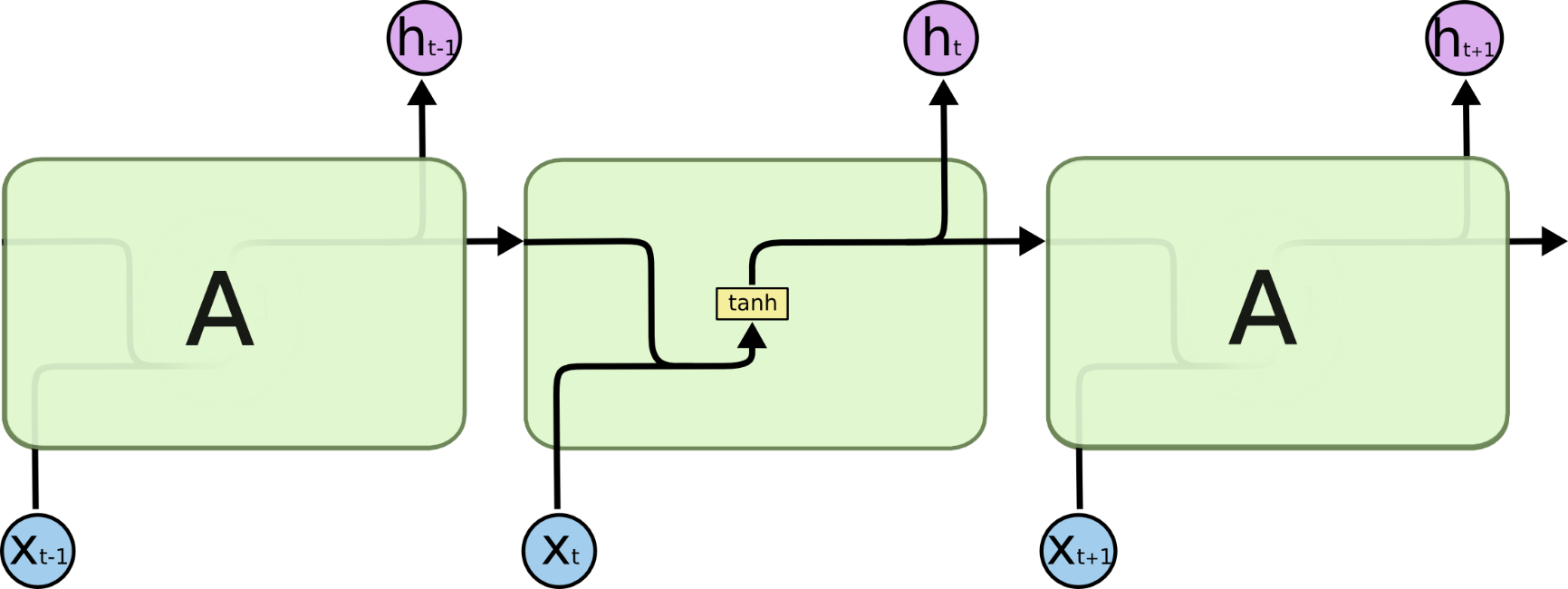
만약 비트코인이 1%가 오른다면 100만원이 1% 이므로 1만원 즉 원금대비 100%의 수익이 발생하는 것이다. 즉 비트코인에서의 선물거래는 레버리지를 일으킨 거래로 이해하면 될 것 같다.



출처 https://cobak.co.kr/community/16/post/107954

**LSTM이란?**

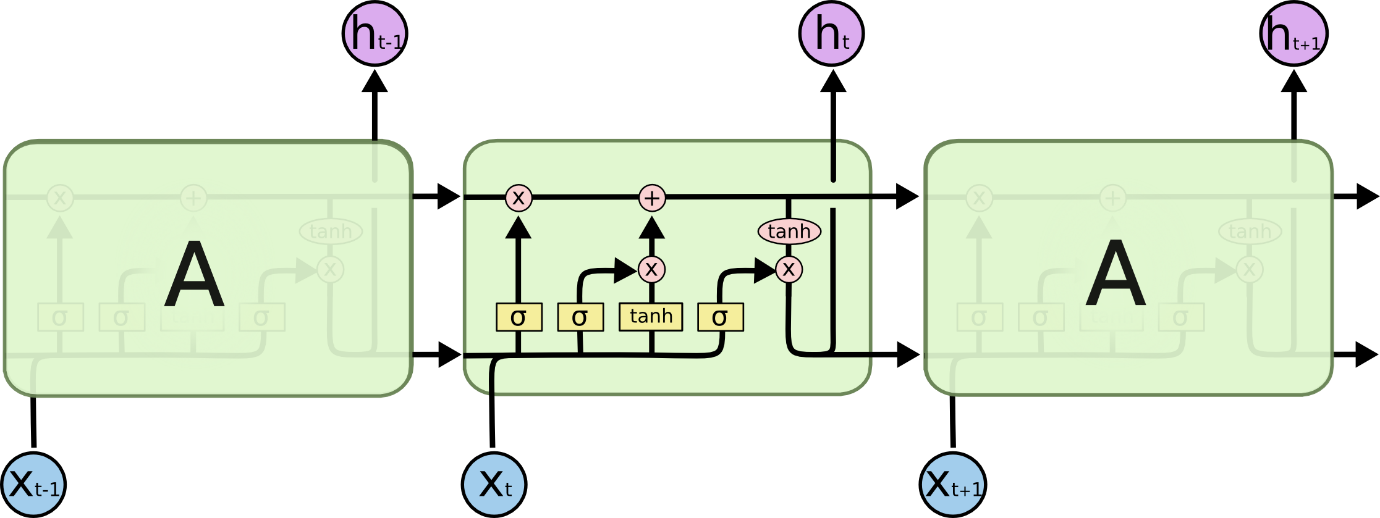
먼저 LSTM을 이해하기 위해서는 앞선 개념인 RNN에 대한 지식이 필요하다. RNN은 시계열 데이터에서 복잡한 패턴을 해석하는데 사용되는 구조이다.



(RNN의 구조)

입력데이터 (Xt)는 시간 축에서 순차적으로 공급해주고 이러한 입력 데이터와 함께 이전 데이터들을 함께 사용하여 새로운 정보를 생성한다. 그리고 이를 출력하는 것이 RNN이다. 그러나 RNN은 장거리 의존성이라는 문제가 있다. 이러한 문제를 이해하기 위해서 예를 들어 보자면. 만약 “The cloud in the sky.”라는 문장이 있다고 가정하자. RNN에 이 문장을 학습시키면 “The cloud in the” 다음에 올 단어는 당연히 “sky”라고 예측할 수 있을 것이다. 그러나 만약 문장이 “I grew up in France ... I speak fluent French”일 때 RNN은 French 관련 정보인 France와의 거리가 멀어 다음에 올 단어들을 예측할 수가 없다. 이것이 장거리 의존성 문제이다.

이러한 문제를 해결하기 위해서 기존 RNN을 수정한 모델인 LSTM(Long short-term memory)이다. LSTM은 기존의 RNN모듈에서 hidden state를 수정하고 cell state(맨 위에있는 줄)를 추가한 구조이다. Cell state는 약간의 변화를 받으며 통과하고 아래의 4개의 레이어(노랑색 상자들)들이 특별한 방법으로 정보를 주고받으면서 통과한다.



(LSTM의 구조)

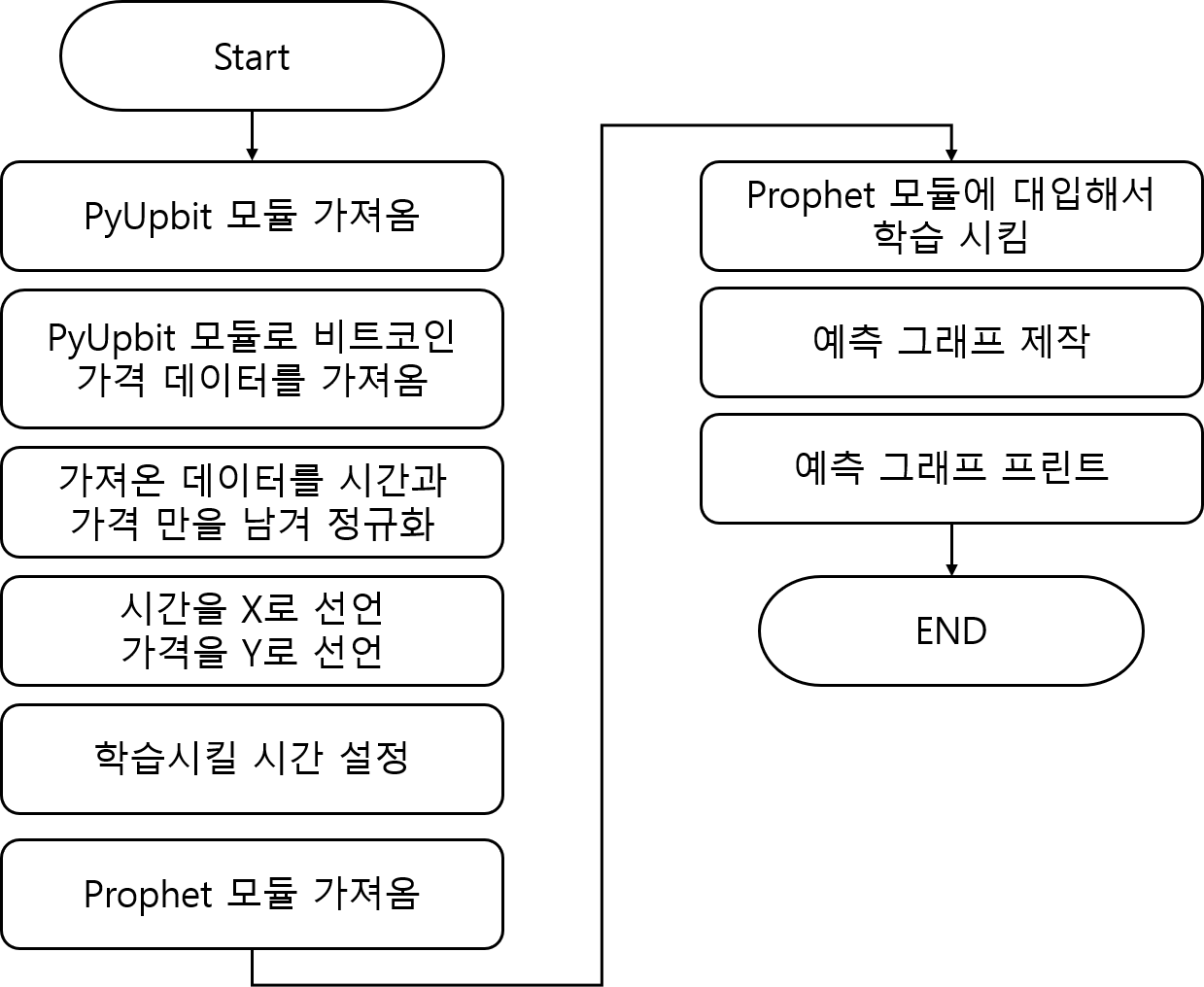
이렇게 레이어들을 통해 필수적이고 중요한 데이터만 남기고 이것을 게이트를 통해 Cell state에 넘긴다. 이렇게 LSTM은 최종적인 Output을 만들어 낸다.

이러한 LSTM을 사용한 모델이 Prophet이다. Prophet이란 Facebook의 Core Data Science팀에서 만든 오픈소스 소프트웨어로 시계열 데이터를 예측하는데 쓰인다. Python[[3]](#footnote-3)에서 사용할 수 있는 일종의 LSTM 모듈이다. 특징으로는 본래 바탕이 되는 ARIMA라는 모델과는 다르게 모델을 차분화해서 정규화 시킬 필요도 없고 [[4]](#footnote-4)결측치 들을 만들어 넣을 이유가 없다. 또 한 회귀분석과 매우 유사하여 생소한 시계열분석에 사용자가 빠르게 적응할 수 있다는 특징이 있다.

**3.비트코인 데이터 예측 모델 구축 및 실험**

3-1 알고리즘 구축 및 구현

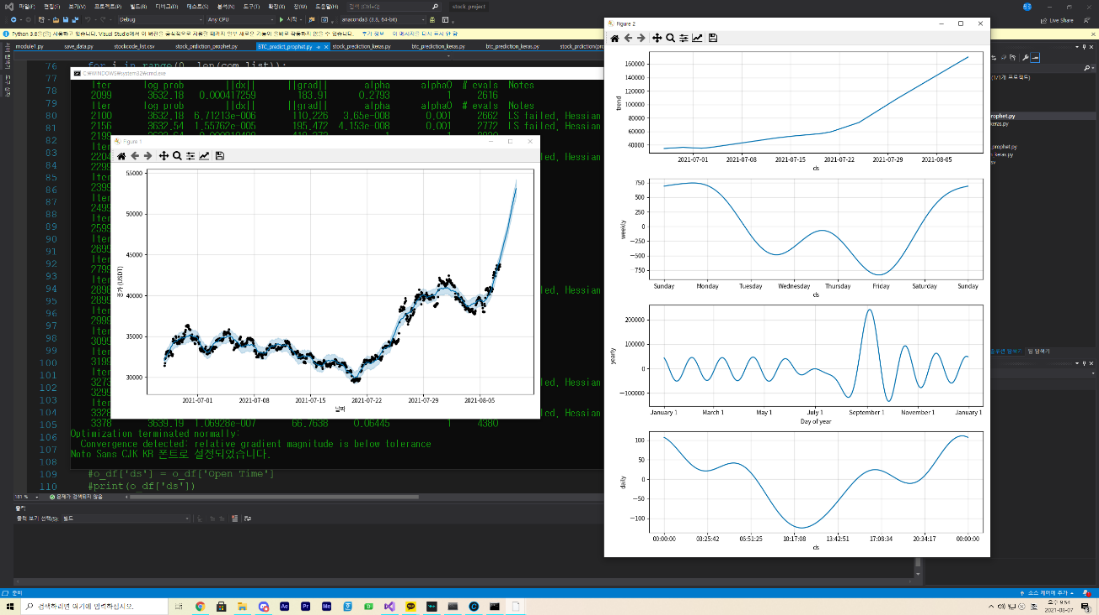
일단 코딩을 먼저하기전에 데이터의 흐름을 이해하기 위해 알고리즘을 먼저 그렸다.



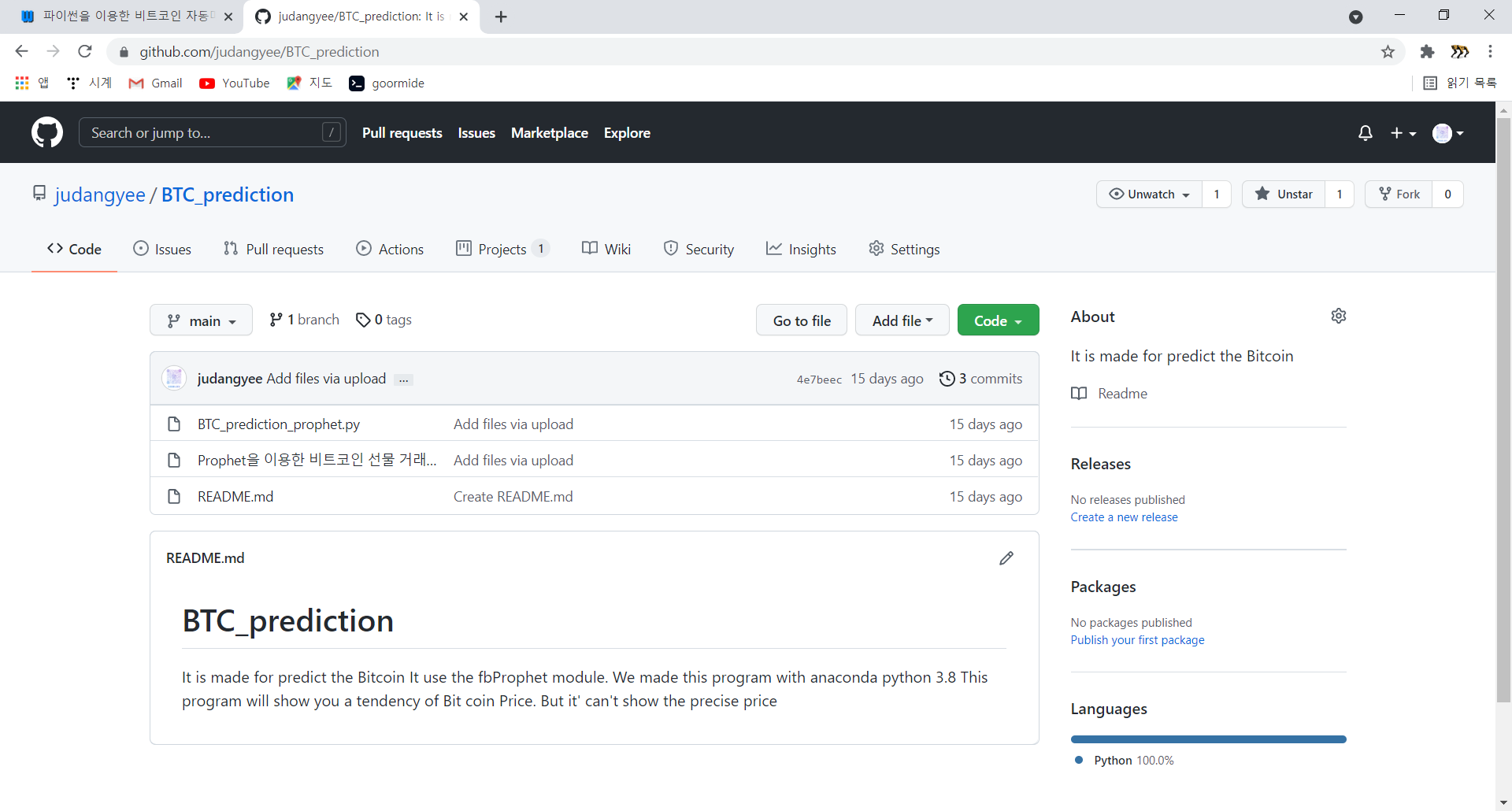
먼저 Pyupbit 모듈로 비트코인의 가격데이터를 가져온다. 그리고 데이터를 가격과 시간만 남겨서 정규화 시키고 시간을 X라는 변수로 선언하고 가격은 Y라는 변수로 선언한다. 이는 X,Y로 그래프를 그리기 위함이다. 다음으로 학습시킬 시간을 정하고 그 시간만큼 데이터를 모듈에 학습 시킨다. 이때 학습하는 과정에서 비정상성 데이터를 제거해야 한다. 그러니 이 모듈은 스스로 비정상성 데이터를 찾아 제거하기에 알고리즘 순서도에서는 제외했다. 그리고 이를 바탕으로 모듈을 활용해 예측 그래프를 그려 낸다. 이러한 과정들을 반복한다.

3-2 데이터 수집 및 처리

본 연구에 앞서 비트코인의 데이터를 수집하기 위해서 pyupbit[[5]](#footnote-5)라는 모듈을 사용해서 데이터를 수집했다. 2000개의 데이터를 3분봉으로 가져왔고 원래의 데이터는 최고가, 최저가, 12시를 기준으로 한 종가를 가져왔고 데이터를 종가와 시간만을 남겨 정규화 했다. 이를 Prophet모듈로 학습시키고 앞으로의 4시간을 예측했다. 이를 바탕으로 전체적인 기울기를 바탕으로 앞으로의 포지션을 잡았고 저점일 때 매수를 하고 최고점에 매도를 하는 방식으로 매매를 진행하였다



<직접 코딩한 화면>



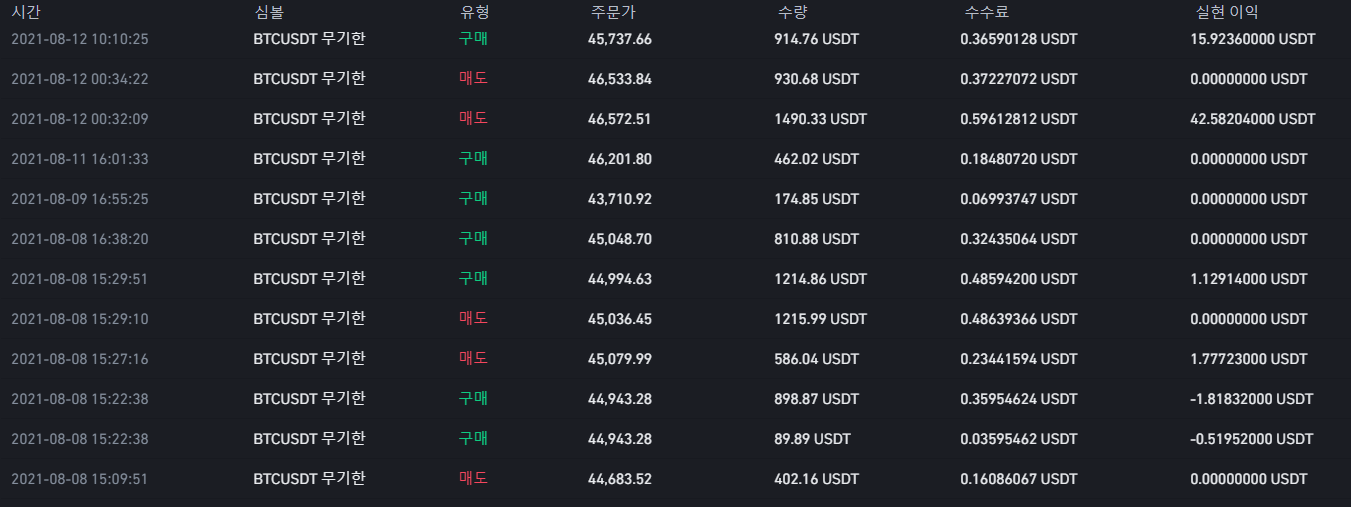
개발자의 깃헙에 코드를 구동 시켜 볼수 있게 코드를 올려 놓았다<https://github.com/judangyee/BTC\_prediction>

3-3 매매 기록 및 설명

매매는 2021년 8월 12일부터 2021년 8월 16일까지 총 5일간 진행했다.



초반에는 우리가 만든 LSTM전략이 맞아 들어가 대략 35%의 수익이 났다.



그러나 거래량을 예측하지 못하는 인공지능 모델의 탓에 예상과는 다르게 비트코인이 급락하기 시작했다. 이때 총잔고가 20만원이 되었다

****

이때 선물거래 특성상 손해는 용납이 되지 않기에 결과 적으로 -25%의 손실을 보고 총잔고 15만원으로 매매를 끝을 냈다.

(자세한 매매 기록은 탐구일지 참조)

1. **평가&결론**

예측 결과 정확하지는 않으나 가까운 시간에서는 가격이 거의 비슷하게 떨어지는 모습을 보였으나 갑자기 급락하거나 급등하는 상황은 예측하지 못했다. 그러나 전체적인 경향성은 맞추는 것을 보 여주었다. 이에는 외부요인들이 작용하는데. 주로 거래량, 언론의 역할이 있다. 이를 위해서 언론, 거래량, 보조 지표들을 분석하는 것이 필요할 것 같다. 또 코인과 같은 변동성이 큰 것들 보다 주식과 같은 변동성이 작은 것에 적용하는 것이 더 효율적일 것 같다. 또한 이러한 모델을 짧은 시간에 수익을 내는 단기적 매매에 사용하기에는 전체적인 범위만을 표시 해주기에 적합하지 않다고 판단할 수 있었다.

1. **실험 환경**

하드웨어 실험 환경은 CPU는 Intel i7 6700k, GPU는 GTX980Ti, DDR4메모리 500GB Samsung evo pro 500GB, 윈도우10, 파이썬 버전은 Anaconda python 3.8.7의 스펙을 가진 컴퓨터에 노트북 컴퓨터로 원격 연결하여 코딩하고 실행시켰다

<참고문헌>

[1] Joo, I.-T., & Choi, S.-H. (2018). 양방향 LSTM 순환신경망 기반 주가예측모델. 한국정보전자통신기술학회논문지, 11(2), 204–208. <https://doi.org/10.17661/JKIIECT.2018.11.2.204>

[2] Binance(<https://www.binance.com/en/trade/BTC_BUSD>)

[3] 업비트(https://upbit.com/exchange?code=CRIX.UPBIT.KRW-BTC)

[4] Bamberg E., 김상욱. "LSTM 인공신경망을 이용한 중국 주가 예측" 동북아경제연구 32, no.2 (2020): 61-84.

[5] 머신러닝 딥러닝 기초 |주|윌텍

[6] 파이썬을 이용한 비트코인 자동매매(개정판) 지은이: 조대표 외 1명

인공지능을 이용한 비트코인 매매 탐구일지

Bitcoin Trading with Artificial Intelligent

연구자

성복고등학교(10618) 송주영

1. LSTM: 장단기 메모리 순환 신경망(뒤에서 자세히 설명함) [↑](#footnote-ref-1)
2. 블록체인: 소규모 데이터들을 분산형으로 저장해 여러 대의 컴퓨터에 저장하는 분산형 저장 기술로 여러 대의 컴퓨터가 기록을 검증해 해킹을 막는 기술이다. [↑](#footnote-ref-2)
3. Python: 코딩 언어 중 하나, 귀도 반 로섬이 만들었다 [↑](#footnote-ref-3)
4. 결측치: 간단히 말해 필요 없는 값 [↑](#footnote-ref-4)
5. Pyupbit: Upbit의 저자들이 파이썬에서 API를 쓰기 쉽도록 만든 모듈 [↑](#footnote-ref-5)