

딥러닝 모델과 통계적 모델을 활용한 비트코인 시세 분석 및 예측

유재준, 홍주오, 채진원, 박종석, 윤주현

요약 최근 암호화폐거래소로 투자자들이 몰리면서 비트코인 가격이 급등락하고 있다. 본 연구의 목적은 딥러닝 모델과 통계적 모델을 이용하여 비트코인의 가격을 예측하는 것이다. 본 연구에서 사용하는 딥러닝 모델에는 비선형성과 장기기억 특성을 보이는 LSTM모델과 LSTM 모델에 비해 더 간단한 구조를 가지고 있는 GRU모델을 선정하였다. 통계적 모델로는 어떤 시계열에도 적용이 가능하며 특히 시간의 흐름에 따라 자료의 변동이 빠를 때 민감하게 반영하는 ARIMA모델과 정확도가 높고 빠르며 직관적인 파라미터로 모델 수정이 용이하고 페이스북에서 공개한 시계열 예측 라이브러리인 Prophet 모델을 선정하였다. 본 연구에서는 FianceDataReader 라이브러리를 사용하여 2018년 1월 1일부터 2022년 12월 31일 데이터를 수집하고 분석하는데 사용하였다.

주제어 : 암호화폐거래소, 비트코인, 딥러닝, LSTM, GRU, ARIMA, Prophet

Abstract Recently, the price of Bitcoin has been fluctuating rapidly as investors flock to cryptocurrency exchanges. The purpose of this study is to predict the price of Bitcoin using a deep learning model and a statistical model, and to analyze whether or not Bitcoin is profitable through an investment strategy. For the deep learning model used in this study, the LSTM model, which exhibits nonlinearity and long-term memory characteristics, and the GRU model, which has a simpler structure than the LSTM model, were selected. As a statistical model, it can be applied to any time series. In particular, the ARIMA model, which sensitively reflects when data changes rapidly over time, and the model can be easily modified with high accuracy, fast, and intuitive parameters, and time series prediction published by Facebook. The library Prophet model was selected. In this study, the FianceDataReader library was used to collect and analyze data from January 1, 2018 to December 31, 2022.

key words : Cryptocurrency Exchange, Bitcoin, Deep Learning, LSTM, GRU, ARIMA, Prophet

논문 작성기간 : 2023년 03월 02일 ~ 2023년 06월 20일

추가적인 연구기간 : 2025년 08월 2일 ~ 2025년 08월 20일

목 차

제1장. 서론	3
1. 연구 배경 및 목적	3
제2장. 이론적 배경	4
1. 암호화폐 시장 및 비트코인의 개념과 각 모델에 대한 설명	4
1) 암호화폐 시장 및 비트코인의 개념	4
2) 딥러닝 모델과 통계적 모델에 대한 설명	5
제3장. 본론	10
1. 데이터 설명	10
2. 일론 머스크 발언으로 인한 각 모델에 대한 분석	10
3. 우크라이나 전쟁으로 인한 각 모델에 대한 분석	13
제4장. 실험 결과 분석	16
1. 각 모델에 대한 실험결과 분석	16
제5장. 결론	17
1. 각 모델에 대한 결론	17
참고문헌	18

1. 서론

비트코인은 전 세계적으로 다양한 암호화폐거래소에서 거래가 이루어지고 있으며 국내에서도 빗썸 거래소(bithumb)를 비롯한 여러 거래소가 암호화폐 투자자들을 끌어들이고 있다. 비트코인을 비롯한 수백 종류의 암호화폐는 하루 거래금액이 6조원대를 돌파하면서 주식시장과 맞먹는 크기로 성장하고 있다. 가격이 급등하자 많은 투자자들이 암호화폐 거래에 뛰어들고 있으며, 가격의 급등락으로 오히려 큰 손해를 보고 있는 투자자들도 나타나면서 사회 문제화되기도 하였다. 암호화폐거래소의 구조는 증권거래소와 비슷하며, 비트코인을 비롯한 암호화폐 가격들은 주가보다 더 변동성이 크고 방향성이 강한 추세적 움직임을 보여주고 있다. 투자 전략적 관점에서 비트코인과 같은 암호화폐는 방향성과 변동성 측면에서 전통적인 주식이나 선물상품보다 더 매력적인 투자 대상이라고 할 수 있다. 주가 예측 방법으로는 기본적 분석이나 기술적 분석이 활용되고 있다. 그러나 주가의 비선형성, 장기 추세, 순환 변동이나 무작위 교란 같은 불안정 요인들로 인해 전통적인 접근 방법들은 예측력이 현저히 떨어지는 한계점이 지적되고 있다. 이에 따라 통계적 접근 방법을 대체하는 다양한 인공지능기법들이 연구되고 있다.

암호화폐 시장이 확대됨에 따라 주가 예측모형을 활용한 암호화폐 가격 예측 연구들이 활발히 이루어지고 있다. 전통적 통계모형인 ARIMA, VAR, ECM 모형 등을 활용한 암호화폐 가격 예측 연구가 대표적이다. 하지만 본 연구에서는 다른 연구와는 다르게 딥러닝모델(LSTM, GRU)과 통계적모델(ARIMA, PROPHET)을 각각 비교 및 분석 할 것이다.

본 연구의 목적은 LSTM을 이용하여 비트코인의 가격을 예측하고, 예측된 가격을 이용하여 비트코인의 투자전략을 제안하며, 실제 비트코인 가격 자료를 이용하여 실증 분석을 통해 제안된 투자전략의 수익성을 분석하는 것이다. 또한 각 모델별로 임의의 사건 A, B로 나누어서 비트코인 시세예측 프로그램에 맞는 모델을 선정하는 것에 목적이 있다. 비트코인과 같은 암호화폐는 주식과는 달리 내재가치의 판단이 불가능하다. 그러나 대부분의 암호화폐 투자자들은 마치 주식 투자하듯 투기적 거래에 뛰어들고 있다. 내재가치를 판단할 수 없는 상황에서는 기본적 분석보다는 본 연구에서 제안하는 모형과 같은 기술적 분석 접근법이 유효하다.

본 연구는 기존의 연구와는 다음과 같은 차이점이 있다. 임의로 사건 A, B로 나누었다. 사건 A는 2021년 5월 13일 일론 머스크가 “테슬라의 비트코인 결제를 잠정 중단하겠다”라고 말한 사건이고, 사건 B는 2022년 2월 24일 우크라이나와 러시아 전쟁을 한 사건이다. 각 모델별로 사건 A, B에 대한 사건 발발후 비트코인 시세 분석 및 예측하여 그래프로 시각화하는 차이점을 가졌다.

2. 이론적 배경

2-1) 암호화폐 시장 및 비트코인의 개념

Satoshi Nakamoto는 2009년 최초의 암호화폐 비트코인을 개발하였다. 비트코인은 디지털 화폐로서 제안되었으며 뒤를 이어 이더리움 등 다양한 형태의 암호화폐들이 등장하고 있다. 최근에는 영국이나 일본 정부처럼 비트코인을 공식 화폐로 인정하려는 분위기가 감지되고 있지만, 대부분의 투자자들은 비트코인을 투자의 수단으로 인식하고 비트코인 자체의 거래에 집중하고 있는 상황이다. 미국의 선물거래소는 비트코인선물을 상장시켜 비트코인을 투자 대상으로 더욱 확장하였다. 국내에서도 비트코인을 비롯한 많은 암호화폐가 여러 거래소를 통해 주식시장과 같이 활발하게 거래되고 있다. 국내 가상화폐거래소에서 거래되고 있는 암호화폐의 수도 200개가 넘고 있으며 비트코인의 시가총액도 900조 원을 넘어서며 삼성전자의 시가총액 500조 원보다도 크다.

2017년 전 세계에서 가장 큰 관심을 끌었던 비트코인은 2008년 말 사토시 나카모토의 논문 한 편으로 세상에 공개되었다. 비트코인 피자 날로 불리는 5월 22일은 최초로 비트코인을 이용한 오프라인 결제가 이루어진 날이다. 2010년 5월 18일에 비트코인 포럼에서 1만 비트코인(그 당시 39달러)으로 피자 두 판을 구매하기를 원했고 나흘 뒤 포럼의 한 유저가 피자 두 판을 구매자의 집으로 보내면서 거래가 성사되었다. 비트코인은 채굴되기도 1년 동안 가치가 0원이었는데 그 이유는 비트코인을 거래하는 거래소가 없었기 때문이다. 2010년 2월 세계에서 최초 비트코인 거래소인 마운트곡스가 생기면서 비트코인이 거래되었고 2010년 비트코인 최고가격은 0.39달러였다.

비트코인 가격의 비선형성과 불확실성 등으로 인해 통계적 방법보다는 기계학습과 같은 인공지능 모형들을 이용한 가격 예측 연구가 활발히 진행되고 있다.

2-2) 딥러닝 모델과 통계적 모델에 대한 설명

(1) LSTM(Long Short Term Memory) 모델

RNN(Recurrent Neural Network)의 한 종류로, 시계열 데이터와 같은 순차적인 데이터를 처리하기 위해 고안된 모델로, RNN(Recurrent Neural Network)의 변형 모델이다.

LSTM은 RNN과 달리 Hidden State와 Cell State라는 두 가지 상태 값을 사용하며, 이를 통해 순차적 데이터에서 정보를 보존, 활용하고, 각각의 Gate에 따라서 어떤 정보를 저장할 것인지, 어떤 정보를 삭제할 것인지, 어떤 정보를 출력할 것인지를 결정한다. LSTM은 순차적인 데이터에서 장기적인 의존 관계를 학습할 수 있어서, 자연어 처리, 음성 인식 등 다양한 분야에서 활용된다.

다음으로는 LSTM 모델의 구성이다.

LSTM은 메모리 셀, 입력 게이트(input gate), 출력 게이트(output gate), 그리고 삭제 게이트(forget gate)라는 4개의 주요 요소로 구성됩니다. 각 요소는 다음과 같은 역할을 한다.

첫 번째로는 메모리 셀이다. 메모리 셀은 현재 시간 단계에서의 상태 값을 저장한다.

두 번째로는 입력 게이트이다. 입력 게이트는 현재 입력을 어느 정도 반영할 지를 결정한다.

세 번째로는 출력 게이트이다. 출력 게이트는 메모리 셀의 값을 현재 상태에 얼마나 반영할 지를 결정한다.

네 번째로는 삭제 게이트이다. 삭제 게이트는 이전 상태에서 어떤 정보를 삭제할 지를 결정한다.

다음으로는 LSTM 모델의 작동원리이다.

첫 번째로는 입력 데이터의 전처리이다. 입력 데이터는 순차적인 시계열 데이터로 이루어져 있다. 따라서 이 데이터를 모델이 학습할 수 있는 형태로 변환해야 한다.

입력 데이터를 일정 길이의 시퀀스로 나누고, 각각의 시퀀스를 벡터 형태로 변환한다.

두 번째로는 LSTM 모델의 학습이다. LSTM 모델은 입력 데이터와 이전 시점의 Hidden State, Cell State를 받아서 다음 Hidden State와 Cell State를 계산한다.

모델은 입력 데이터를 받아 Hidden State와 Cell State를 계산하고, 이를 출력한다. 이러한 과정을 반복하면서 LSTM 모델은 순차적인 데이터의 패턴을 학습한다.

세 번째로는 예측이다. 학습된 모델을 이용하여 새로운 시계열 데이터를 예측한다.

입력 데이터를 이용해 Hidden State와 Cell State를 계산하고, 다음 Hidden State와 Cell State를 예측하고 이를 반복하여 다음 시계열 값을 예측한다.

(2) GRU(Gated Recurrent Unit) 모델

GRU는 LSTM의 장기 의존성 문제에 대한 해결책을 유지하면서, 은닉 상태를 업데이트하는 계산을 줄였다. 다시 말해서, GRU는 성능은 LSTM과 유사하면서 복잡했던 LSTM의 구조를 간단화 시켰다. GRU는 LSTM과 마찬가지로 입력 데이터와 이전 상태를 이용하여 출력을 계산한다. 하지만 LSTM의 cell state와 hidden state 대신에, GRU는 업데이트 게이트와 리셋 게이트 두 가지를 사용합니다. 업데이트 게이트는 이전 상태에서 얼마나 많은 정보를 유지할 것인지를 결정하며, 리셋게이트는 얼마나 많은 이전 상태를 현재 상태에 반영할 것인지를 결정한다.

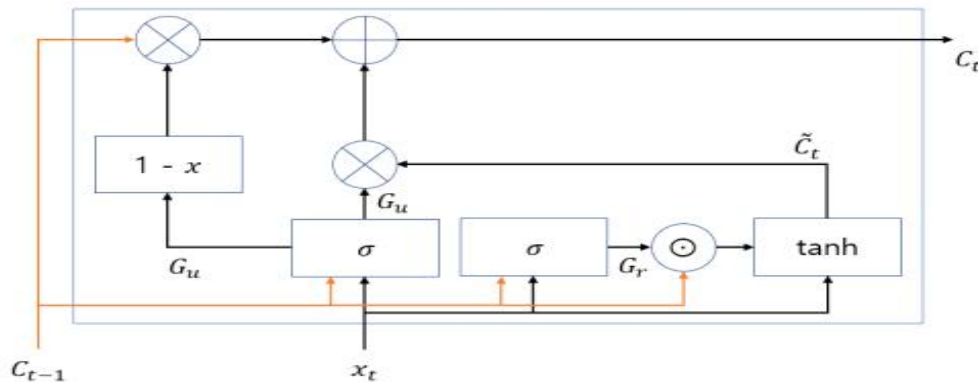
GRU의 구조는 다음과 같다.

입력 게이트는 현재 입력 데이터를 이전 상태와 함께 처리하기 위한 게이트이다.

리셋 게이트는 이전 상태의 얼마나 많은 정보를 유지할 것인지를 결정하기 위한 게이트이다.

새로운 상태는 입력 데이터와 이전 상태, 리셋 게이트를 고려하여 계산된 새로운 상태 값이다. 업데이트 게이트는 이전 상태에서 얼마나 많은 정보를 유지할 것인지를 결정하기 위한 게이트이다. 출력 게이트는 새로운 상태와 업데이트 게이트를 이용하여 계산된 출력 값이다.

다음으로는 GRU 작동원리에 대해 알아보겠다.



GRU 모델은 두 개의 게이트를 이용해 입력을 연산한다.

첫 번째로는 업데이트 게이트이다. 업데이트 게이트는 이전 시점의 출력과 현재 시점의 연산결과 중 어떤것에 더 가중치를 주는 가를 결정해준다.

$$G_u = \sigma(W_u[C_{t-1}, x_t] + b_u).$$

* σ = 시그모이드 활성화 함수

W_u = 업데이트 게이트의 연결계수

C_{t-1} = 이전 시점의 출력값

x_t = 현재 시점의 입력값, b_u = 업데이트 게이트의 바이어스값

두 번째로는 리셋 게이트이다. 리셋게이트는 이전 시점의 출력값을 새로운 현재 시점의 출력이 될 출력 후보값에 얼마나 포함시킬 것인지 결정해준다.

$$G_r = \sigma(W_r[C_{t-1}, x_t] + b_r).$$

* σ = 시그모이드 활성화 함수

W_r, b_r = 각각 리셋게이트의 연결계수와 바이어스값

C_{t-1} = 이전 시점의 출력값

x_t = 현재 시점의 입력값

마지막으로 연산한 결과들을 토대로 현재시점의 출력값을 결정하는 단계이다.

$$C_t = G_u \tilde{C}_t + (1 - G_u) C_{t-1}.$$

* 업데이트 게이트값 : G_u

현재시점에서의 후보값 : \tilde{C}_t

이전시점의 출력값 : C_{t-1}

업데이트 게이트값은 시그모이드 함수의 출력값이기 때문에 출력범위가 0~1사이가 된다.

업데이트 게이트의 값이 0이라면 최종 출력에서 현재시점에서의 후보값이 반영되지않고 이전 시점의 출력값이 그대로 출력된다.

반대로 업데이트 게이트값이 1이라면 이전시점의 출력값은 현재 시점의 출력에 반영되지 않고 현재시점에서의 후보값으로 결정된다.

이러한 순환 연산과정은 입력이 모두 입력되어 연산되고 출력으로 나가는 과정이 끝날때까지 반복된다.

(3) ARIMA(Autoregressive Integrated Moving Average) 모델

ARIMA 모델은 시계열 데이터를 예측하는데에 사용되는 모델 중 하나이다.

ARIMA 모델은 시계열 데이터의 경향성, 패턴, 계절성을 파악하여 이를 기반으로 미래 값을 예측한다.

ARIMA 모델은 다음의 세 가지 요소를 기반으로 예측한다.

AR(Autoregression)은 자기회귀 모형으로, 이전 값이 이후 값을 결정하는 모델입니다. 현재의 관측치가 이전 관측치들의 선형조합으로 이루어진다고 가정한다.

I(Integrated)는 누적된 차분을 통해 시계열 데이터의 경향성을 제거하는 모델이다.

MA(Moving Average)는 이동평균 모형으로, 이전의 오차들의 선형조합으로 이루어진다고 가정한다.

ARIMA 모델의 예측 결과는 시계열 데이터의 예측 구간과 신뢰 구간으로 나타내어진다. 이 구간은 일반적으로 95% 신뢰도로 계산된다.

ARIMA 모델은 시계열 데이터의 분석 및 예측에 널리 사용되며, 특히 경제학, 금융, 생물학 등에서 사용된다.

ARIMA 모델의 작동원리는 다음과 같다.

첫 번째로는 데이터의 정상성 검사이다. ARIMA 모델은 정상성 시계열 데이터에서만 적용 가능합니다. 따라서, 우선 주어진 데이터가 정상성을 가지는지 검사한다.

두 번째로는 모형 설정이다. ARIMA 모델은 AR(Auto Regressive), MA(Moving Average), 및 차분(Difference)의 3가지 파라미터를 사용하여 모델을 설정한다.

AR(Auto Regressive) 파라미터는 과거의 데이터를 이용하여 현재를 예측하는데 사용된다. 즉, 이전 시점의 자료가 현재의 자료에 영향을 주는 정도를 나타내는 파라미터이다. MA(Moving Average) 파라미터는 백색잡음(white noise)을 이용하여 현재 값을 예측하는데 사용된다. 백색잡음은 평균이 0이고 분산이 일정한 무작위한 값이다. 차분(Difference) 파라미터는 시계열 데이터가 비정상성을 가지는 경우, 차분을 통해 정상성 시계열 데이터로 변환하여 모델링한다.

세 번째로는 모수 추정이다. 모델의 파라미터를 추정한다. ARIMA 모델은 최적의 파라미터를 추정하기 위해 ACF(Auto-Correlation Function)와 PACF(Partial Auto-Correlation Function) 그래프를 이용한다.

네 번째로는 모형 검토이다. 추정된 모델이 데이터를 얼마나 잘 예측하는지 검토한다. 예측된 모델의 잔차(Residual)는 백색잡음(white noise)이어야 한다. 만약 잔차가 백색잡음이 아니면, 모형이 좋지 않은 것이다.

다섯 번째로는 모형 예측이다. 추정된 모델을 사용하여 시계열 데이터를 예측한다. ARIMA 모델은 향후 n 개의 값을 예측할 수 있다.

ARIMA 모델은 시계열 데이터에서 규칙성을 찾아 예측하는데 유용한 모델이다. 그러나, 모델을 적용하기 전에 데이터의 정상성을 검사하는 것이 중요하다. 또한, 모델의 파라미터를 추정하고 모형 검토를 신중히 수행하여 좋은 예측 결과를 얻을 수 있다.

(4) PROPHET 모델

PROPHET 모델은 Facebook에서 개발한 시계열 예측 모델로, 간단하고 유연한 방법으로 시간에 따른 데이터의 패턴을 모델링하고 예측할 수 있다. 이 모델은 특히 계절성 패턴을 갖는 데이터에 적합하다.

PROPHET 모델의 작동원리는 다음과 같다.

첫 번째로는 트렌드 구성이다. PROPHET 모델은 데이터에 존재하는 주기성, 성장 또는 감소하는 추세 등과 같은 트렌드를 모델링한다. 이를 위해 시계열 데이터를 주간, 월간 또는 연간의 트렌드 구성 요소로 분해한다.

두 번째로는 계절성 모델링이다. PROPHET은 주기적인 계절성 요소를 모델링하여 데이터의 계절성 변동을 포착합니다. 계절성은 연간, 월간, 주간 등 다양한 주기로 발생할 수 있으며, PROPHET은 이를 감지하고 예측에 반영한다.

세 번째로는 휴일 효과 고려이다. PROPHET은 휴일과 같은 이벤트가 데이터에 미치는 영향을 고려한다. 휴일은 예측에 미치는 영향이 크므로, PROPHET은 휴일의 효과를 모델에 포함하여 더 정확한 예측을 가능하게 한다.

네 번째로는 유연한 모델링이다. PROPHET 모델은 다양한 데이터 패턴을 수용할 수 있는 유연한 모델이다. 예측 대상 데이터에 대한 여러 가정을 자동으로 설정하고, 이를 조정할 수 있는 매개 변수를 제공하여 모델을 더욱 유연하게 조정할 수 있다.

다섯 번째로는 예측 결과 해석이다. PROPHET은 트렌드, 계절성, 휴일 효과 등을 고려하여 예측 결과를 제공한다. 이를 통해 시계열 데이터의 미래 동향을 예측할 수 있으며, 결과를 해석하고 활용할 수 있다.

PROPHET 모델은 비교적 간단하고 사용하기 쉽지만, 다양한 시계열 데이터에 대해 강력한 예측 성능을 발휘할 수 있다. 그러나 모델의 성능은 데이터의 특성과 사용자의 설정에 따라 달라질 수 있으므로, 적절한 모델 설정과 조정이 중요하다.

3. 본론

3-1) 데이터 소개

본 연구에서는 FinanceDataReader 라이브러리를 사용하였다.

FinanceDataReader 라이브러리는 한국 주식 가격, 미국주식 가격, 지수, 환율, 암호화폐 가격, 종목 리스팅 등 금융 데이터 수집 라이브러리이다. 또한 금융 데이터를 다루는데 가장 기본이 되는 데이터는 거래소별 전체 종목 코드와 가격 데이터이다.

FinanceDataReader 라이브러리에는 종목 코드에는 KRX (KOSPI, KODAQ, KONEX), NASDAQ, NYSE, AMEX, S&P 500이 있고, 가격 데이터에는 해외주식가격데이터(애플, 아마존 등), 국내 주식 가격 데이터(삼성전자, 셀트리온헬스케어 등), 각종 지수(코스피 지수, 코스닥 지수, 나스닥지수), 암호 화폐가격(BTC/USD, BTC/KRW), 환율 데이터(원달러환율, 달러당 유로화 환율) 이 있다.

FinanceDataReader 라이브러리에서 2018년 1월1일부터 2022년 12월 31일까지의 데이터를 가져와 비트코인 시세 분석 및 예측에 필요한 데이터를 사용하였다.

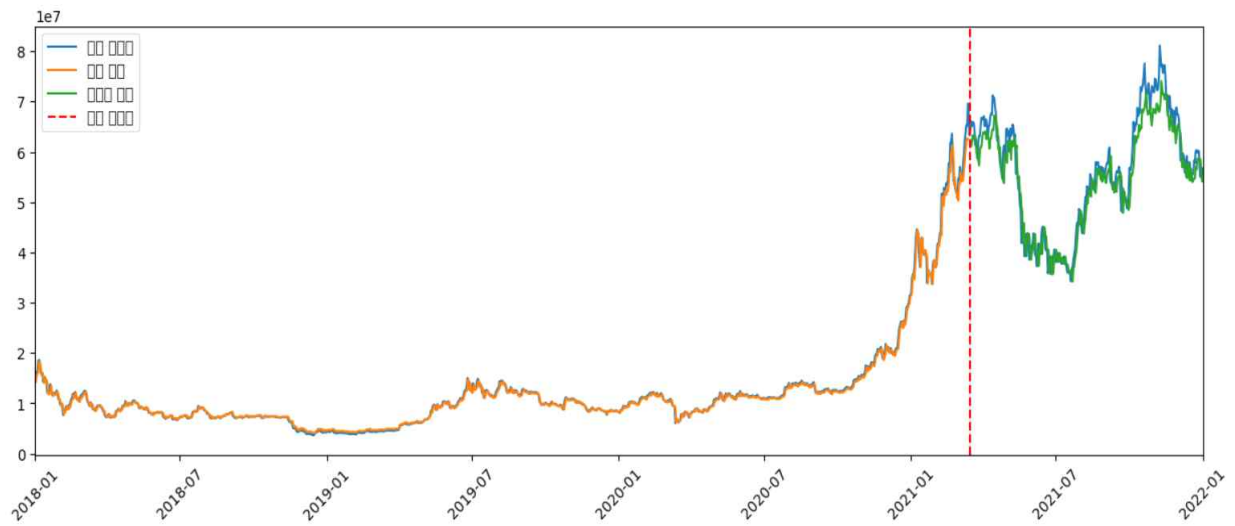
추가적인 연구기간에는 2018년 01월 01일부터 2025년 08월 20일 까지의 비트코인 종가 데이터를 활용하였다.

3-2) 일론 머스크 발언으로 인한 각 모델에 대한 분석

일론 머스크 사건은 2021년 5월 13일이 되자 머스크는 비트코인이 과도한 전력소모와 화석 연료 사용으로 인해 환경 파괴적이라며 테슬라의 비트코인 결제를 잠정 중단하겠다고 밝혔다. 해당 트윗 직후 전체 가상화폐 시가총액 415조원이 하루만에 증발했으며 머스크는 깨진 하트 이모티콘 트윗을 쓰는 등 비트코인과의 이별을 암시하기도 했다.

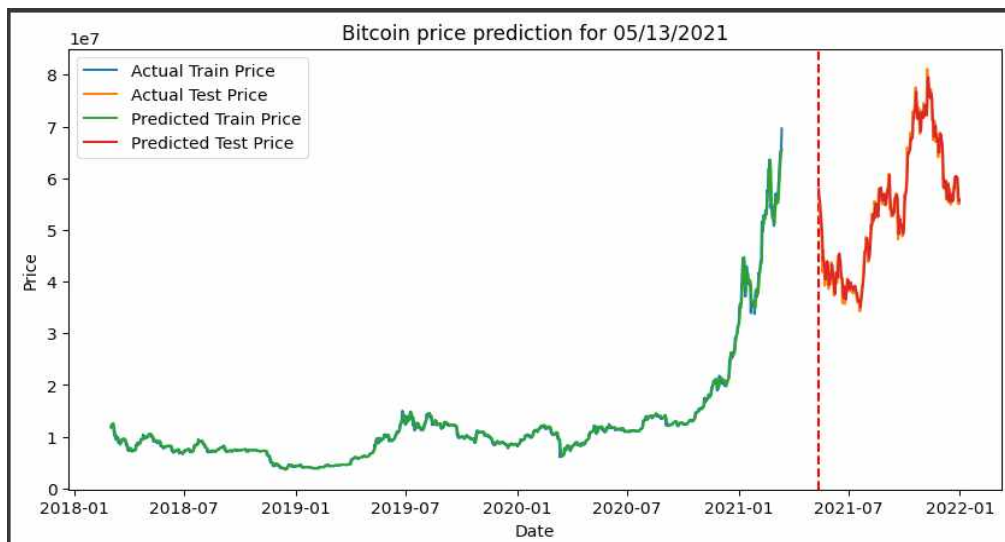
이 때문에 1분기 실적 발표 이후 테슬라가 초고점에서 비트코인을 전량매도했을 것이라는 공포로 패닉셀이 시작되었으나 머스크가 테슬라는 1분기 매도 이후로 한 번도 판매하지 않았다는 트윗을 써 단기적으로 가격이 상승하였다.

(1) LSTM 모델



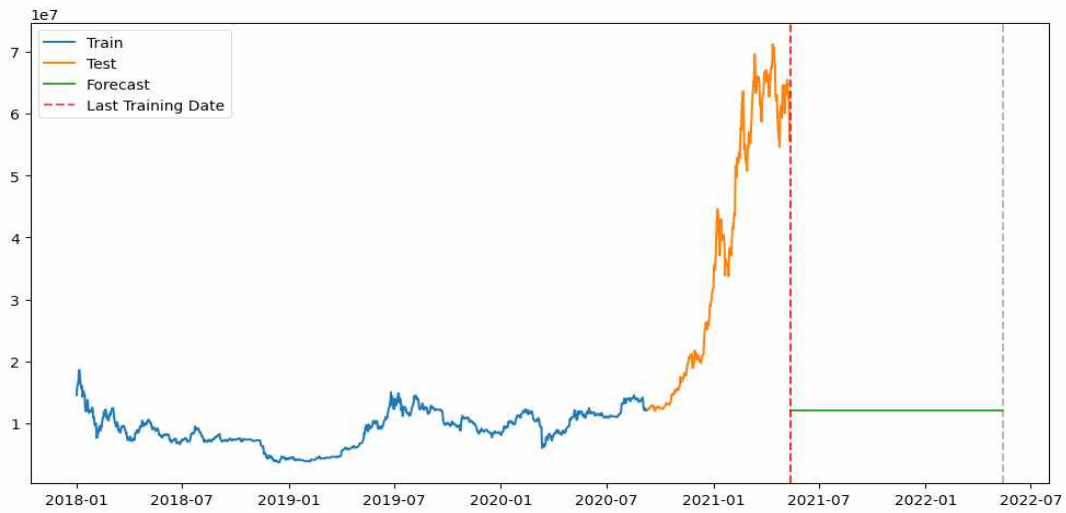
LSTM 모델인 경우 사건이 발발후 비트코인 시세가 상승하였지만 얼마 시간이 안가 비트코인 시세가 확 떨어지는 것을 볼 수 있다.

(2) GRU 모델



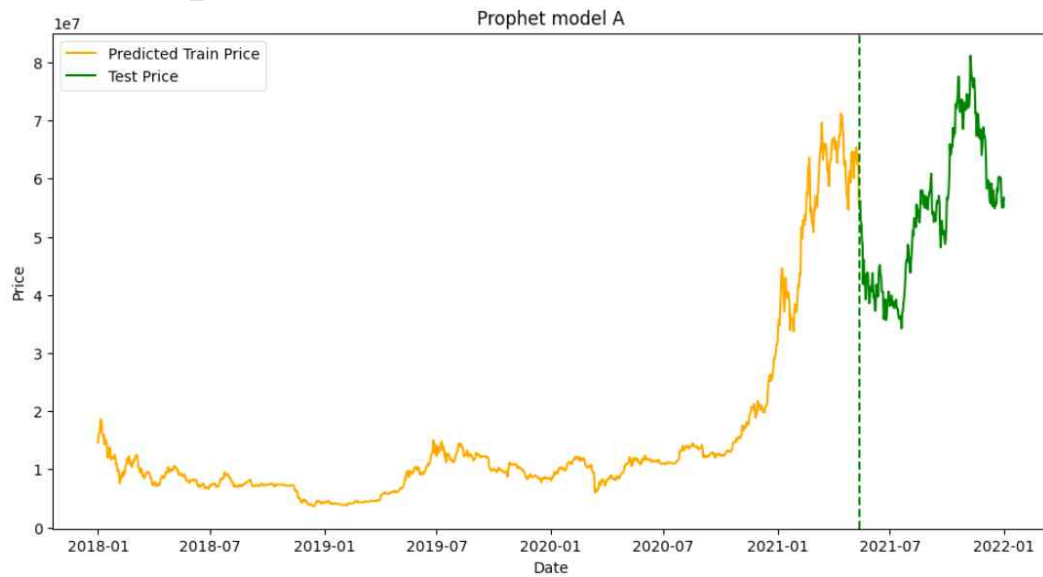
GRU 모델인 경우 사건이 발발후 바로 비트코인 시세가 떨어지고 약 2달후 다시 오르는 것을 볼 수 있다.

(3) ARIMA 모델



ARIMA 모델은 사건이후 예측이 되지않는 것을 볼 수 있다.

(4) PROPHET 모델



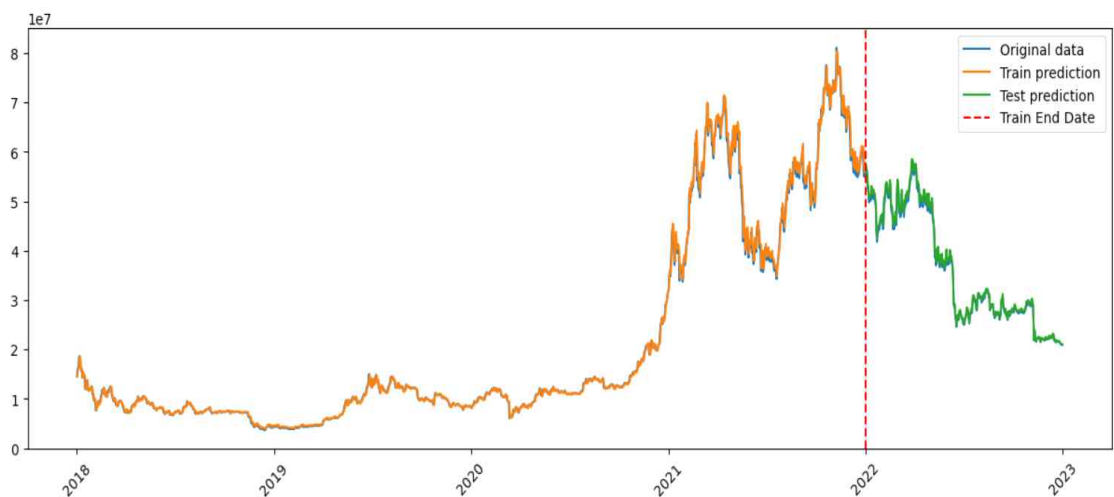
PROPHET 모델인 경우 사건이 발발후 바로 떨어지는 것을 볼수 있고, 얼마지나지 않아 다시 코인시세가 올라가는 것을 볼 수 있다.

3-3) 우크라이나 전쟁으로 인한 각 모델에 대한 분석

우크라이나 전쟁은 2022년 2월 24일 러시아 대통령 블라디미르 푸틴이 특별 군사작전 개시 명령을 선언한 이후 러시아가 우크라이나를 침공하면서 발발한 전쟁이자 21세기 2번째로 일어난 유럽 국가 간의 전쟁이다.

푸틴은 우크라이나의 비무장화, 비나치화, 돈바스주민 보호를 명분으로 내세우며 전쟁을 개시하였다.

(1) LSTM 모델



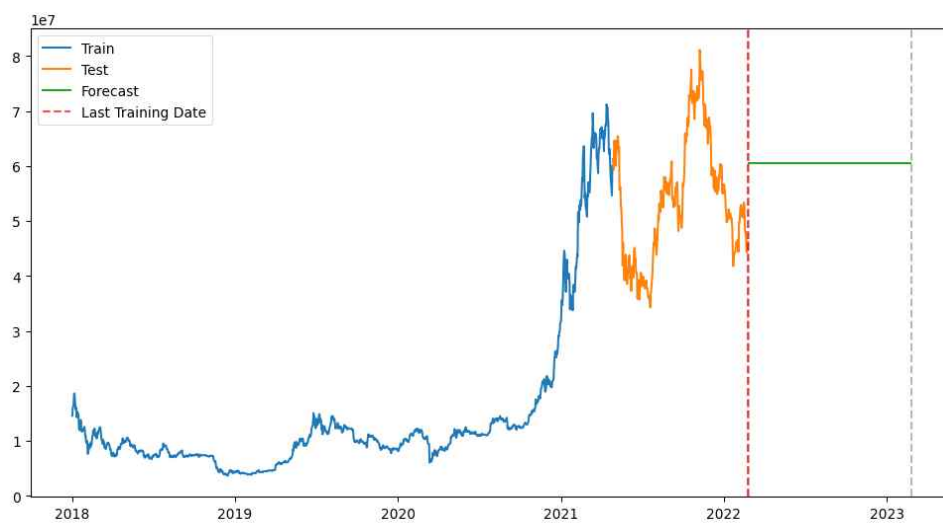
LSTM 모델인 경우 우크라이나 전쟁 발발이후 떨어지다가 2022년 중반으로 갈수록 눈에 띄게 떨어지는 것을 볼 수 있다.

(2) GRU 모델



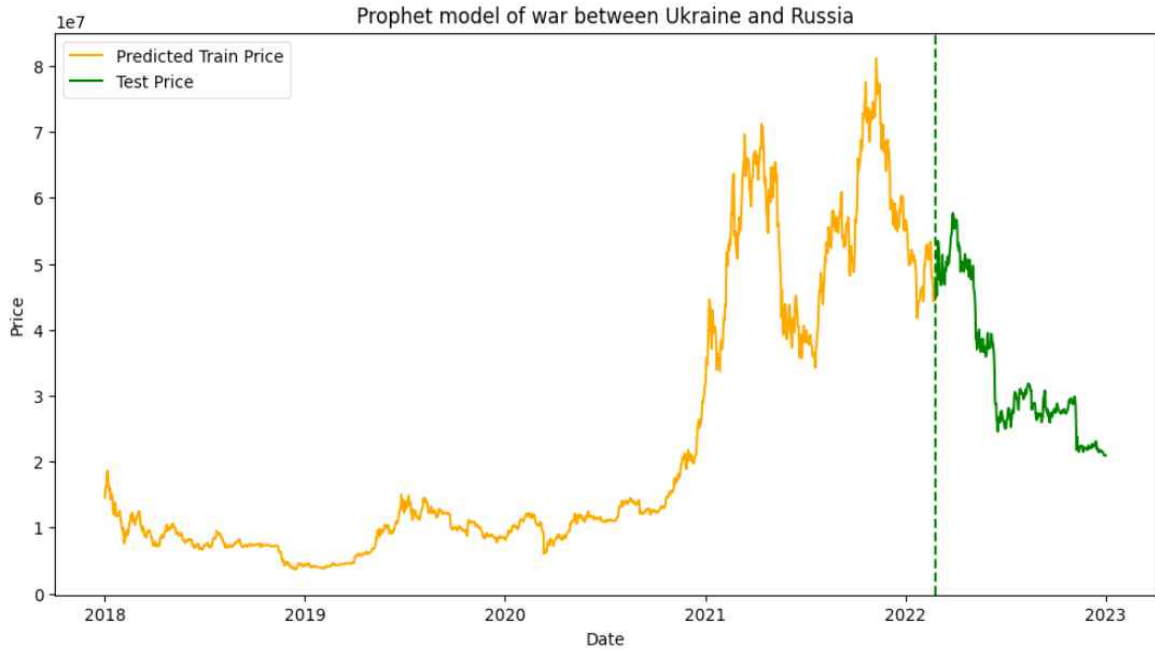
GRU모델인 경우 우크라이나 전쟁 발발이후 단기적으로 상승하였다가 떨어지는 것을 볼 수 있다.

(3) ARIMA 모델



ARIMA 모델은 사건이후 예측이 되지않는 것을 볼 수 있다.

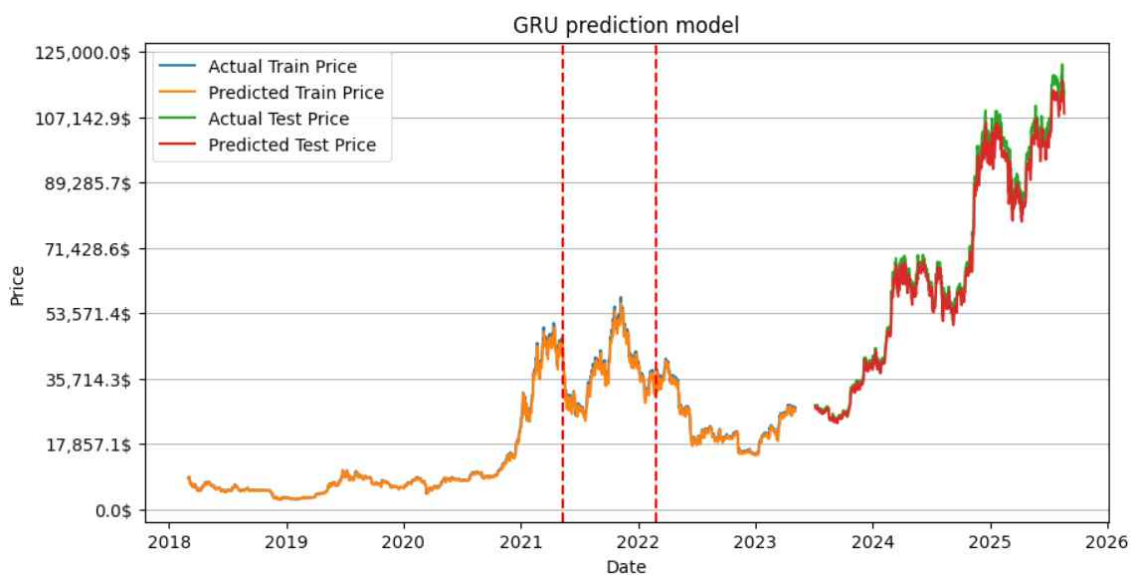
(4) PROPHET 모델



PROPHET 모델인 경우 전쟁발발 이후 미세하게 상승하였다가 떨어지는 것을 볼 수 있다.

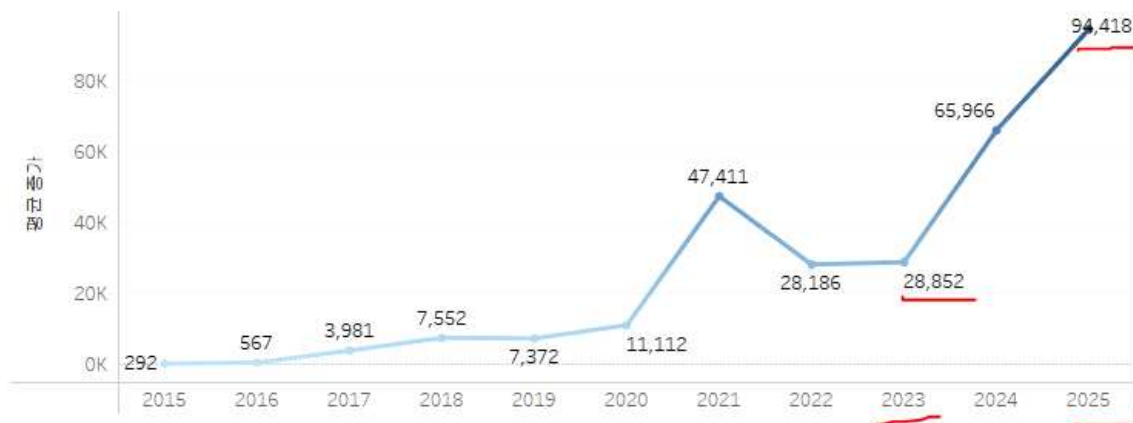
● 추가적인 연구

추가적인 연구기간 : 2025년 08월 2일 ~ 2025년 08월 20일



2018년 01월 01일부터 2025년 08월 20일 까지의 비트코인 종가 데이터를 훈련하여 만든 GRU 모델이다. 가운데 빨간색 선은 앞서 정의를 내린 “우크라이나 전쟁” 과 “일론 머스크발언” 발생 후 비트코인 종가가 어떻게 변화하는지 알아보고자 하였다.
두 사건 모두 떨어졌다가 반등하는 것을 볼 수 있다.

다음으로는 2015년부터 2024년까지의 평균종가 그래프가 담겨져 있는 의미를 알아보고자 한다.



종가가 2023년~2025년 사이에 평균종가가 7만가까이 올랐다.

이 이유를 살펴보니 3가지 이유가 있었다.

첫 번째 반감기로 인한 희소성 증가와 기관 매수세가 맞물려 가격 상승 가능성을 높인다. 비트코인은 4년마다 반감기 시기를 갖고 있다.

두 번째 금리인하 기대감, 정부 및 기관 투자자의 수요증가 하였다.

세 번째 비트코인 현물 ETF 상장승인 하여 누구나 비트코인 투자가 가능 하다.

현물 ETF는 비트코인을 직접보유가 가능하다는 특징이 있고,

선물 ETP는 비트코인 선물계약 이라는 특징이 있다.

4. 실험결과 분석

1) 첫 번째 사건인 일론 머스크 사건에서 각 모델별로 실험결과를 나타내었다.

(1) LSTM 모델에서는 일론 머스크 사건 정의에서 설명하였듯이 사건 발발후 일시적으로 비트코인 시세는 증가하였다.

(2) GRU 모델에서는 사건 발발 후 떨어졌다. 하지만 2021년 7월 이후 증가하는 추세를 보이긴 하였다.

(3) ARIMA 모델에서는 사건 발발후 예측하지 못한 그래프를 보여준다.

(4) PROPHET 모델에서는 GRU 모델과 비슷하게 2021년 7월 이후 증가하는 추세이다.

일론 머스크 사건으로 봤을 때 비트코인 시세예측 프로그램에 적합한 모델은 LSTM 모델이었다.

2) 두 번째 사건인 우크라이나와 러시아 전쟁을 각 모델별로 실험결과를 나타내었다.

(1) LSTM 모델에서는 사건 발발 직후 떨어지고 가면갈수록 떨어지는 것이 눈에 띄었다.

(2) GRU 모델에서는 사건 발발 이후 올라가는 추세를 보이긴 하였으나 몇 달이 지나지 않고 떨어지는 모습을 볼 수 있었다.

(3) ARIMA 모델에서는 사건 이후 예측하지 못하는 것을 보여준다.

(4) PROPHET 모델에서는 GRU 모델과 비슷하게 사건 이후 올라가는 추세를 보이긴 하였으나 얼마 못가 떨어지는 모습을 볼 수 있었다.

날짜별 모델 간의 성능 비교표				
모델 데이터 날짜	LSTM 모델	GRU 모델	ARIMA 모델	Prophet 모델
2018~2021년	96%	97%	-2.3%	97%
2018~2022년	99%	96%	-3.3%	96%
2018~2025년		99%		

위 표는 딥러닝모델과 통계적모델의 날짜별 성능을 비교한 표이다.

전체적으로 성능이 잘 나왔지만 ARIMA 모델의 성능이 떨어진 것을 확인할 수 있다.

4. 결론

본 연구는 딥러닝 모델(LSTM, GRU) 과 통계적 모델(ARIMA, PROPHET)을 일론머스크 발언 사건 및 우크라이나 전쟁에 빗대어 사건 이후 예측하는 방향으로 나아갔다. 사건 이후 예측이 안되는 ARIMA 모델을 제외하고 대체적으로 비슷한 양상이 보여졌다.

본 연구에서 그치지 않고 더 나아가 LSTM, GRU, ARIMA, PROPHET 모델을 제외한 딥러닝 모델과 통계적 모델을 각각 1개 혹은 2개의 모델을 선정한 후 앞서 설명한 방식 말고 다른 방식으로 접근할 것이다. 또한 ARIMA의 정확도가 음의 결정계수로 표현되었다. 그러므로 ARIMA 모델 대신에 Boosting, Gradient 모델로 시도할 것이다. 다른 모델 대신에 더 많은 데이터를 사용하여 정확도를 높일 것이다.

또한 예측 정확도를 높이기 위해 다음과 같은 것을 추가할 것이다.

첫 번째 경제 및 시장 지표(환율, 나스닥 지수) 등 경제지표 추가할 것이다.

경제 지표가 있어야 비트코인 시세와 비교가 될것이라고 생각한다.

두 번째 뉴스 및 소셜 미디어 감성 분석할 것이다.

뉴스가 긍정적이면 1로 부정적이면 0으로 라벨링하여 학습한후 모델링하면 예측 정확도가 높아질것이라고 생각한다.

세 번째 데이터 증강 기법을 활용하여 4년마다 반감기 활용할 것이다.

앞서 언급했듯이 비트코인은 4년에 한 번씩 반감기를 갖는다. 랜덤삭제(RD)를 사용하면 계속 증가하는 것을 막을것이라고 생각한다.

네 번째 종가와 저가를 학습하여 종가 예측할 것이다.

위 3가지 기술을 접목시켜 예측할 것이다.

본 연구에서 언급한 딥러닝 모델은 다른 분야에서도 활용이 된다. 예를 들어 경제 예측을 할 수 있으며 경제 환경에서 통계적 모델은 수요와 공급을 예측하기에 적합하며 그 외에 마케팅 분석을 할때 통계적 모델은 마케팅 분야에서도 활용하기 유용하다고 생각한다. 의료 진단과 의료 분야에서 통계적 모델은 질병 진단, 예후 예측 등에 사용될 수 있을 것으로 보이며 환자의 건강 데이터를 분석하여 특정 질병의 발생 가능성이나 치료 결과를 예측할 수 있을 것으로 예측된다. 더 나아가서 교통 및 수송 계획에 도움이 될것으로 보이는데 그 이유는 교통 및 수송 분야에서 통계적 모델은 교통 흐름, 교통 체증, 이동 패턴 등을 분석하는 데 유용하기 때문이다.

이 외에도 통계적 모델은 금융 분석, 환경 모니터링, 인구 통계학 등 다양한 분야에서 사용될수 있다. 그리고 딥러닝 모델은 음성 인식 시술에 널리 사용될수있으며, 이미지 인식과 자연어 처리, 추천시스템으로 음악, 도서, 영화 등 다양한 추천 시스템에서 딥러닝 모델이 사용되기에 충분한 역량을 갖고 있다.

참고문헌

최희식 (경민대학교 IT경영과) , 조양현 (삼육대학교 컴퓨터학부)

한국멀티미디어학회지 v.19 no.1 ,pp. 28 - 36 , 2015 , 1229-778x , 한국멀티미디어학회

박도형(2018.06.25.)(일반논문.국민대학교)JAKO201820540191717.pdf (koreascience.kr)

양희성 (충남대학교 공과대학 한국전자거래학회지 = The Journal of Society for e-Business Studies v.24 no.4 ,pp. 109 - 118 , 2019 , 2288-3908 , 한국전자거래학회

황용일 (Dept. of Economics, Dongduk Women's University)

한국전자거래학회지 = The Journal of Society for e-Business Studies v.23 no.3 ,pp. 13 - 32 , 2018 , 2288-3908 , 한국전자거래학회

손종혁 (Department of Business Administration, Sangmyung University) , 김정연

(Department of Business Administration, Sangmyung University)

한국전자거래학회지 = The Journal of Society for e-Business Studies v.24 no.4 ,pp. 109 - 118 , 2019 , 2288-3908 , 한국전자거래학회

양희성 (충남대학교 공과대학 정보통신공학과) , 권영미 (충남대학교 공과대학 정보통신공학과)

한국멀티미디어학회지 v.19 no.1 ,pp. 28 - 36 , 2015 , 1229-778x , 한국멀티미디어학회

김선웅.(2021). 딥러닝을 이용한 비트코인 투자전략의 성과분석. 12(4).249-258

전주용,여은정.(2014). 비트코인의 이해: 금융경제학적 관점에서.18(4).211-239

서윤범. "비트코인 가격 등락 예측을 위한 딥러닝 모델 연구." 국내석사학위논문 단국대학교 대학원, 2018. 경기도