

박훈석 정민관 이 현
신영섭 2025.01.09

딥러닝 기반 가상화폐 시세 변동 예측 시스템

사랑합니
다



목차

01

프로젝트 팀
구성 및 역할

02

프로젝트
개요

03

프로젝트 수행
절차 및 방법

04

결론

05

소감



CHAPTER 06

프로젝트 일정





01

프로젝트 팀 구성 및 역할

프로젝트 팀 구성 및 역할

박훈석



- 프로젝트 관리(PM)
- 도메인 분석

정민관



- 데이터 베이스 설계 및 관리
- 업비트 API 연동 및 관리
- 발표

이 현



- 데이터 베이스 설계 및 관리
- 딥러닝 모델링
- PPT 제작

신영섭



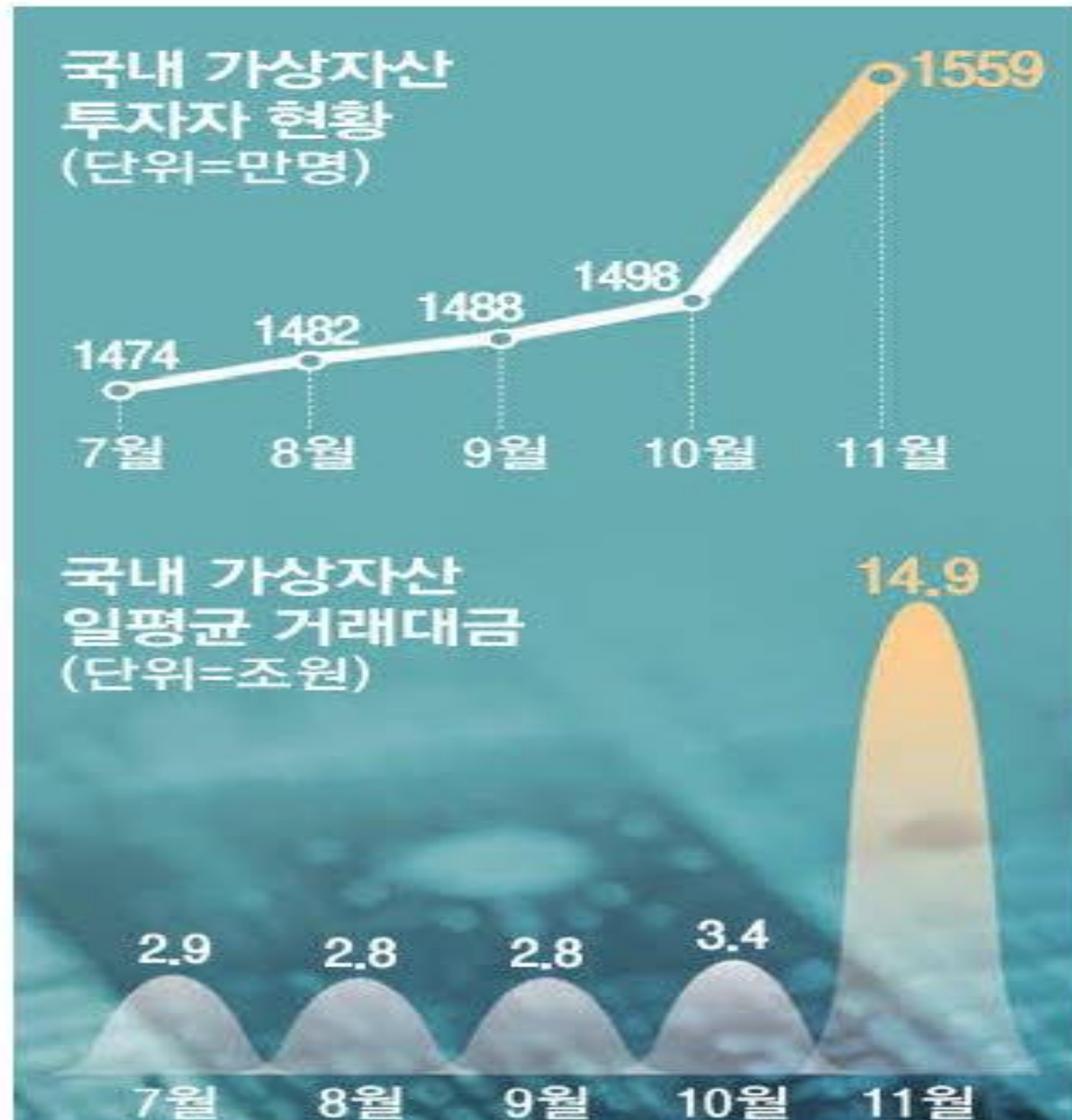
- 데이터 분석 및 전처리
- 딥러닝 모델링
- 시각화



02

프로젝트 개요

가상화폐 시장 현황



미국의 비트코인 ETF

파이낸셜뉴스

구독신청 | fnC

블록체인 > 정책

어엿한 주류금융... 비트코인 ETF 운용자산 1000억弗 돌파

김미희 기자

파이낸셜뉴스 입력 2024.12.30 18:04 수정 2024.12.30 18:04

가+ 가- ㉞

美 비트코인 현물 ETF 승인 1년
블랙록 IBIT 3개월간 40% 수익
가상자산 친화적인 트럼프 2기
기관투자자 시장참여 유인할 듯

주요 비트코인 현물 상장지수펀드(ETF) 현황

(단위 : 달러)

※ 2024년 12월 30일 기준

자료 : 비트보(BITBO)



2024년 가파른 성장을 보인 미국의 비트코인 ETF

- 운용자산(AUM): 1,060억 달러 돌파
- 블랙록 IBIT ETF 수익률: 3개월 간 40% 상승
- 미국 금융시장 편입 비트코인: 113만 개

대한민국의 비트코인 ETF 현황

금융 CEO 절반 "가상자산 투자 허용되면 매수할 것"

서형교 기자 ☆ 조미현 기자 ☆

입력 2024.04.18 18:33 수정 2024.04.19 02:28

☆ ↗ 💬 😊 🖨

은행·증권·운용사 상위 15개사 설문

"韓도 코인 ETF 상장" 만장일치
"비트코인은 가치저장 수단이자
화폐처럼 교환매개 성격도 있어"

법인 가상자산 투자 금지한 한국
"개인들 투기판으로 전략" 지적도

서유석 금투협회장 "가상자산 ETF 추진"

신민경 기자 ☆

입력 2025.01.02 09:42 수정 2025.01.02 09:42

가가

☆ ↗ 💬 😊 🖨

"종투자 해외진출 확대 지원"
BDC 제도 국회 통과도 지원
"주니어 ISA 제도 도입 추진"

그렇다면 대한민국은?

•금융투자협회

- 전략토큰증권 법제화 추진
- 가상자산 ETF 도입 계획
- 디지털 자산시장 인프라 혁신

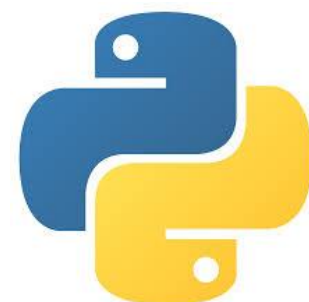
•금융권 인식 변화

- CEO 15인 설문 결과 비트코인 = 가치저장 수단 (6/15명)
- 비트코인 현물 ETF 도입 찬성 (14/15명)
- 법인 가상자산 투자 허용 찬성 (8/15명)

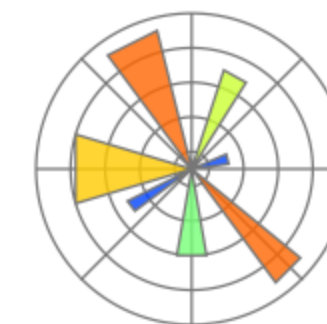
- 금융권, 가상자산에 점차 우호적 태도
- 제도적 기반 마련 필요성 공감

비트코인 연도별 종가 추이





Pandas



03 프로젝트 수행 절차 및 방법

수집 → 전처리 → 모델학습 → 실제예측 적용

기본 컬럼

데이터셋 소개

업비트 api

- 기간: 2017년 9월 ~ 2024년 12월 31일
- 데이터 타입: 일봉 시계열 데이터
- 정렬 방식: 날짜 순 정렬 완료

OHCLV

Opening_price: 시가

High_price: 최고가

Closing_price: 종가

Low_price: 최저가

Volume: 총 거래량

파생 컬럼

이동평균(Moving Average)

단순 이동 평균 (SMA)

- 모든 가격에 동일한 가중치 부여
- 장점: 계산 간단
- 단점: 최근 가격 변화 반영 미흡

가중 이동 평균 (WMA)

- 최근 가격에 높은 가중치
- 최근 데이터의 중요성 강조
- 오래된 데이터에 낮은 가중치 부여

가중이동평균 WMA 사용, 5일
wma_5



주요 특징

- 최근 가격이 다음 기간 예측에 더 큰 영향
- 데이터의 시간적 가중치 차등화
- 시장 트렌드 민감도 향상

파생 컬럼

볼린저 밴드(Bollinger Band)

주가의 변동에 따라 상하밴드의 폭이 같이 움직이게 하여 주가의 움직임을 밴드 내에서 판단하고자 고안된 주가지표

중단밴드 - middle_band, $MBB=SMA(n)$

상단밴드 - upper_band, $UBB=MBB+(\sigma[TP,n] \times m)$

하단밴드 - lower_band, $LBB=MBB-(\sigma[TP,n] \times m)$

• $\sigma[TP,n]$: n 기간 동안의 표준편차
• TP :

대표가격 (고가+저가+종가) ÷ 3 (고가+저가+종가) ÷ 3

• m : 표준편차 배수 (일반적으로 2)

$\%B = (\text{종가} - \text{하단 밴드}) / (\text{상단 밴드} - \text{하단 밴드})$ 밴드내 현재 위치

$\text{BandWidth} = (\text{상단 밴드} - \text{하단 밴드}) / \text{중간 밴드 밴드 대역폭}$

파생 컬럼

볼린저 밴드(Bollinger Band)



매수 신호

- 주가 상단 밴드 반복 터치/돌파
- 밴드 상방 확장
- 가격 중간선 상단 움직임
- RSI 상승 모멘텀 확인

매도 신호

- 주가 하단 밴드 지속 터치
- 밴드 하방 수축
- 가격 중간선 하단 움직임
- 거래량 하방 돌파 확인

파생 컬럼

MACD(Moving Average Convergence Divergence, 이동평균 수렴확산 지수)

주가 추세의 강도, 방향, 모멘텀 및 지속 시간의 변화를 나타내도록 설계

12일 EMA에서 26일 EMA를 빼서 계산(단기에서 장기를 뺌)

MACD는 두 EMA의 수렴 및 발산에 관한 것

- 추세 전환점 포착
- 모멘텀 변화 감지
- 트렌드 추종 전략

• $MACD = 12일\ EMA - 26일\ EMA$

•신호선 = MACD 라인의 9일 EMA

•히스토그램 = MACD라인 - 신호선

12일 EMA: 단기 추세를 나타냄
26일 EMA: 장기 추세를 나타냄
9일 EMA: 신호선 계산에 사용

파생 컬럼

MACD(Moving Average Convergence Divergence, 이동평균 수렴확산 지수)



매수 신호

- MACD 라인 제로선 상향 돌파
- 시그널 라인 위로 상승
- MACD 히스토그램 양(+)의 영역 확장
- 상승 모멘텀 지속적 확인

매도 신호

- MACD 라인 제로선 하향 돌파
- 시그널 라인 아래로 하락
- MACD 히스토그램 음(-)의 영역 확장
- 하락 모멘텀 지속적 확인

파생 컬럼

RSI(Relative Strength Index)

주가의 평균 상승폭과 하락폭을 비교하여 가격의 상승 압력과 하락 압력 간의 상대적인 강도

일반적으로 9일, 14일 또는 25일 RSI를 계산

상대강도(RS) 계산

$RS = \text{평균 상승분} / \text{평균 하락분}$

RSI 최종 공식

$RSI = RS / (1 + RS) \times 100$

파생 컬럼

RSI(Relative Strength Index)



매수 신호

- 평균 이득 > 평균 손실
- RS(Relative Strength) > 1
- RSI 지속적인 상승
- RSI 50선 상향 돌파
- RSI 70 이상으로 상승

매도 신호

- 평균 이득 < 평균 손실
- RS(Relative Strength) < 1
- RSI 지속적인 하락
- RSI 50선 하향 돌파
- RSI 30 이하로 하락

파생 컬럼

Stochastic Oscillator

확률적 오실레이터는 특정 기간 동안 증권의 종가를 가격 범위와 비교하는 모멘텀 지표

Stochastic Oscillator 계산

$$\%K = (C - L_5) / (H_5 - L_5)$$

- H_5 : 지난 5일 최고가
- L_5 : 지난 5일 최저가
- 현재가: 당일 종가

$$\%D = \%K \text{의 3일 SMA}$$

%D 계산

- %K의 3일 단순이동평균(SMA)
- 마지막 3개 %K 값 평균

파생 컬럼

Stochastic Oscillator



매수 신호

- 20 이하로 내려갔다가 상승할 때
- %K가 %D선을 아래에서 위로 돌파 (골든 크로스)
- 50선 상위 유지

매도 신호

- 80 이상 올라갔다가 하락할 때
- %K가 %D선을 위에서 아래로 돌파 (데드 크로스)
- 50선 하위 유지

과매수/과매도 상태 판단

- 0~100 사이 값
- 20 이하: 과매도
- 80 이상: 과매수

파생 컬럼

이동 평균 계산의 데이터 결측 문제

- 결측치 발생 원인

시계열 데이터의 시간순 이동 평균 계산 -> 이동 평균
-> 계산의 데이터 공백 현상 발생

Ex) 5일 이동 평균의 경우 -> 1~4일 데이터 -> 결측치, 5일째 데이터부터 정산
계산

- 해결 방법

결측치 데이터 삭제 -> 완전한 이동평균 확보

종속변수 설정

주가 방향성 예측 : 이진 분류 모델

목표

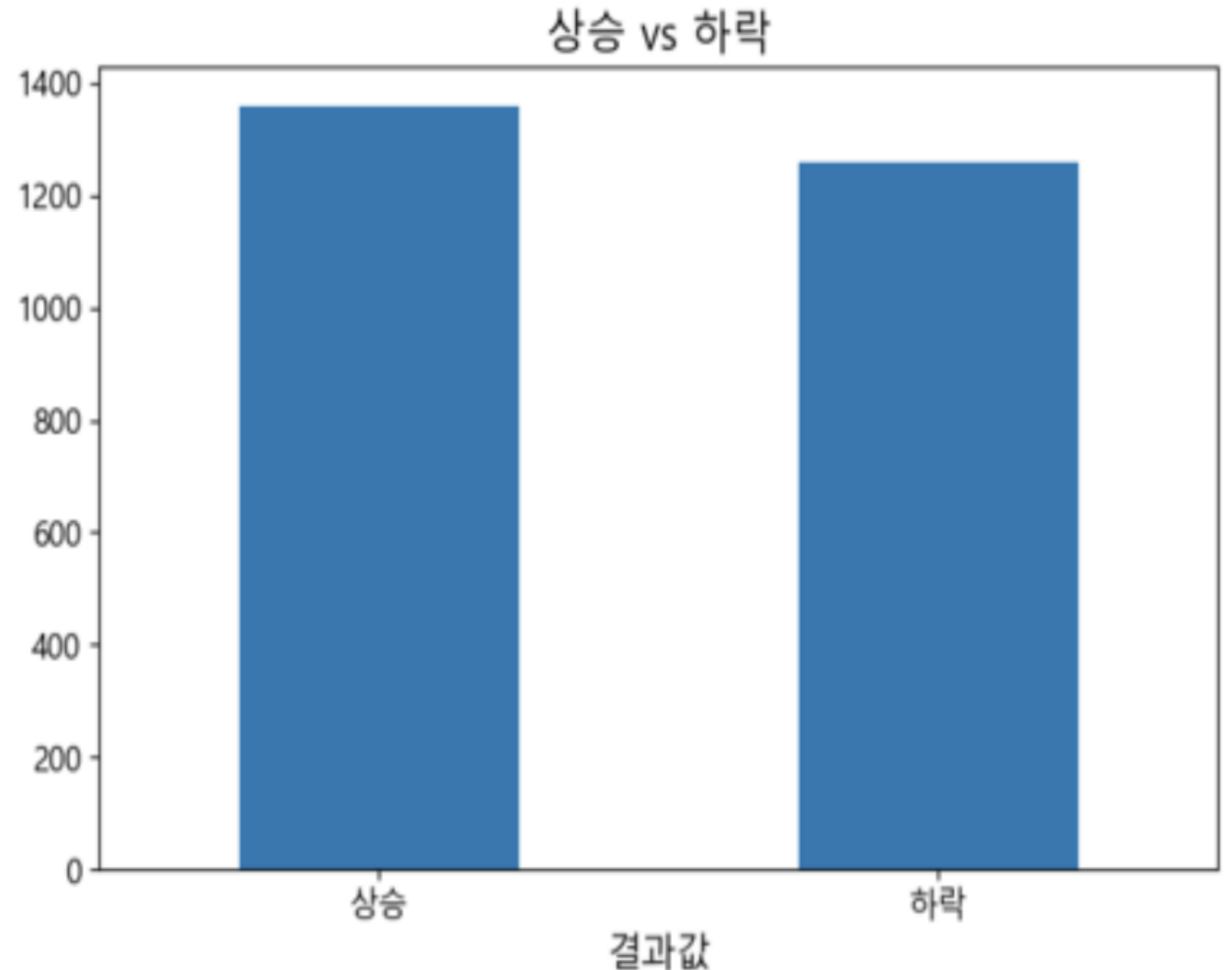
- 다음날 주가 상승/하락 예측
- 1(상승), 0(하락) 이진 분류

데이터 처리 방법

- 현재 종가 vs 다음날 종가

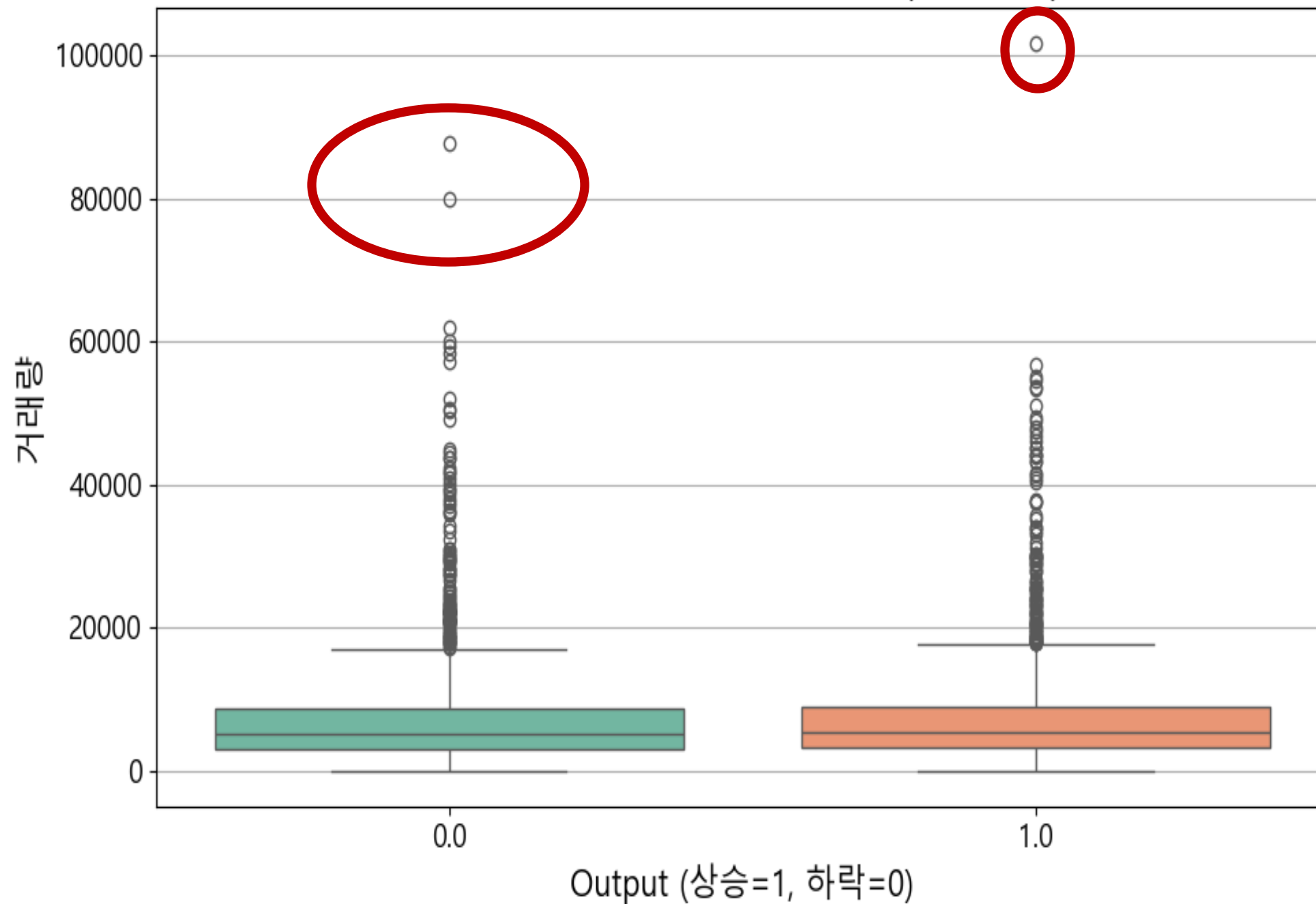
비교

- 상승 시 1, 하락 시 0 라벨링



독립변수 선택

상승, 하락에 따른 거래량 분포 (박스플롯)



상승, 하락에 따른 거래량 분포가 균일하여
거래량 변수는 제외

➔ 이상치 : 2017년 말, 2018년 초
(비트코인 거래가 가장 활발했던 연도)

독립변수 선택

변수명	설명	분류
opening_price	시가	기본 가격 데이터
high_price	고가	기본 가격 데이터
low_price	저가	기본 가격 데이터
closing_price	종가	기본 가격 데이터
wma_5	5일 가중이동평균	이동평균 지표
middle_band	볼린저 밴드 중심선	변동성 지표
upper_band	볼린저 밴드 상단선	변동성 지표
lower_band	볼린저 밴드 하단선	변동성 지표
%B	볼린저 밴드 위치 지표	변동성 지표
bandwidth	볼린저 밴드 폭	변동성 지표
rsi_9	9일 상대강도지수	모멘텀 지표
macd	MACD 선	추세 지표
macd_signal	MACD 신호선	추세 지표
%K	스토캐스틱 %K선	모멘텀 지표
%D	스토캐스틱 %D선	모멘텀 지표

기존 4가지 컬럼 외 다양한 가격정보, 이동평균, 볼린저 밴드, 모멘텀 지표를 포함하여 가격 변동을 예측

-> 가격의 움직임에 대한 다양한 패턴을 학습하고 미래 가격 변동을 예측하는데 더 좋은 성능이 될 것이라고 예상

CHAPTER 3

모델링

DNN

- 여러 개의 레이어로 구성된 인공 신경망. 입력층, 은닉층, 출력층으로 나뉘며 은닉층이 깊은 구조를 가짐.
- 정형데이터, 이미지 분류, 음성 인식 등 다양한 분야에 사용.
- 레이어 별로 활성화 함수와 가중치를 통해 비선형성을 학습.

LSTM

- RNN의 한 종류로, 장기 의존성 문제를 해결하기 위해 고안됨.
 - 일반 RNN보다 더 긴 시퀀스를 학습할 수 있음.
- 시계열 데이터, 자연어 처리, 음성 인식, 주가 예측에 사용

모델 선택

RNN

- 이전 타임스텝의 출력을 현재 타임스텝 입력으로 다시 사용하는 재귀적 구조.
- 시계열 데이터 예측, 자연어 처리, 음성 데이터 처리에 활용.
- 장기 의존성 문제

Bidirectional RNN

- RNN의 확장으로, 데이터를 양방향으로 처리
 - 자연어 처리에서 문맥 정보를 더 풍부하게 활용
- 문장이나 시간적 데이터에서 앞뒤 문맥이 중요한 경우
효과적

모델링

데이터셋 분할

학습 및 테스트: ~2024-6-30

검증: 2024-7-1 ~ 2024-12-4

실제 예측 적용: 2024년 12월 데이터

모델	Accuracy	Precision	Recall	F1 Score
DNN	0.8153	0.9242	0.7176	0.8079
RNN	0.5102	0.5542	0.5679	0.5610
LSTM	0.5510	0.5510	1.0000	0.7105
B-RNN	0.5333	0.5496	0.8675	0.6729

CHAPTER 3

모델링

구분	파라미터	설정값	선택 이유
네트워크 구조	은닉층 수	5	모델 복잡도 확보
	은닉층 노드 수	100	충분한 학습 용량
학습 설정	에포크	100	과적합 방지
	배치 사이즈	10	작은 데이터셋 고려
	학습률	0.001	안정적 학습
최적화	옵티마이저	Adam	효율적 학습
	손실 함수	Cross-Entropy	이진 분류에 적합
	활성화 함수	ReLU	기울기 소실 방지
스케일링	min-max scaler	0~1	가격 데이터의 상대적 위치 파악 용이

주요 특징

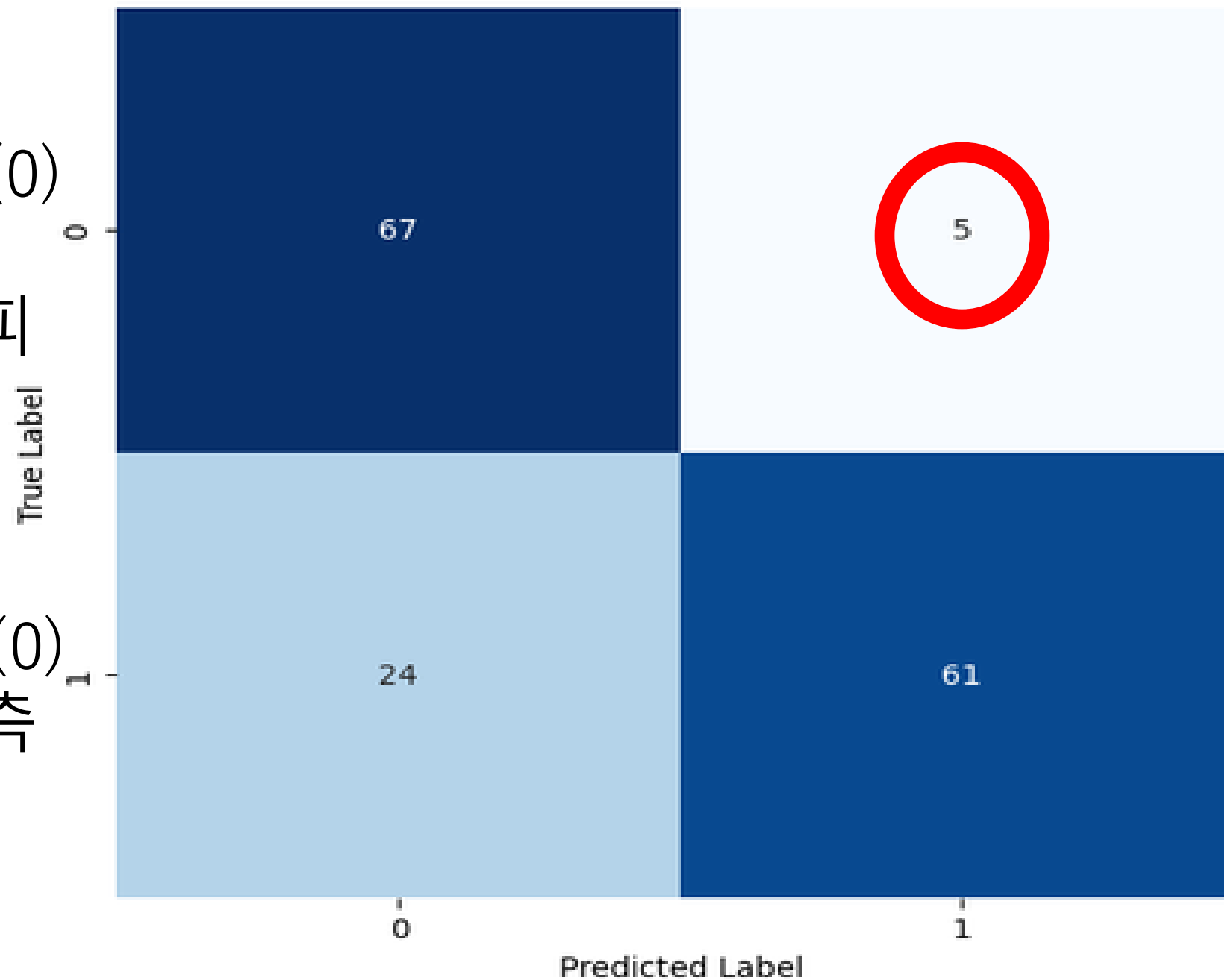
- 과적합 방지를 위한 적절한 에포크와 배치 사이즈 설정
- Adam 옵티마이저로 학습 효율성 확보
- ReLU 활성화 함수로 학습 안정성 향상

설계 의도

- 데이터 부족 문제 고려
- 1(상승), 0(하락) 이진 분류
- 과적합/과소적합 균형 유지

모델링

Validation Confusion Matrix



TN (True Negative)

- 실제: 하락(0), 예측: 하락(0)
- 하락을 정확히 예측
- 매매 미실행으로 손실 회피

FP (False Positive)

- 실제: 하락(0), 예측: 상승(1)
- 잘못된 상승 예측
- 실제 손실 발생 구간

FN (False Negative)

- 실제: 상승(1), 예측: 하락(0)
- 상승을 하락으로 잘못 예측
- 기회비용 발생 구간

TP (True Positive)

- 실제: 상승(1), 예측: 상승(1)
- 상승을 정확히 예측
- 실제 수익 발생 구간

암호화폐 VS 주식

구분	암호화폐	주식
발행 주체	블록체인 기반 분산 발행	기업이 발행하는 지분증권
거래 시간	24시간 365일 거래 가능	장 운영 시간에만 거래 가능
가치 결정	시장 수요와 공급에 따라 결정	기업의 실적과 성과에 기반
규제	최소한의 규제, 높은 변동성	엄격한 규제와 법적 보호
소유권	디지털 자산의 소유	기업의 부분 소유권
수익 구조	시세 차익만 가능	배당금 + 시세 차익
투자 진입 장벽	낮음 (소액 투자 가능)	상대적으로 높음
시장 안정성	높은 변동성	상대적으로 안정적

암호화폐와 주식을 분석하기 위해서는 다른 전략 필요

비트코인 1월 시세 예측

날짜	실제	예측	수익률(종가)
1/1	상승	상승	1.482%
1/2	상승	상승	1.789%
1/3	상승	상승	0.561%
1/4	상승	상승	0.062%
1/5	하락	상승	-0.031%
1/6	상승	상승	3.849%
1/7	하락	상승	-3.643%
1/8	하락	하락	-1.67%

1월 투자 시뮬레이션 결과 분석

1/1: +1.482% (상승)

1/2: +1.789% (상승)

1/3: +0.561% (상승)

1/4: +0.062% (상승)

1/5: -0.031% (하락)

1/6: +3.849% (상승)

1/7: -3.643% (하락)

1/8: -1.670% (하락)

전략 1: 홀드 전략(1/1-1/8)

- 원금 유지
- 총 누적 수익률: +4.069%
- 투자 특징: 변동성 감소

전략 2: 단기 매매 전략

- 최적 매매 시점: 1/1 매수 → 1/6 매도
- 최대 수익률: +2.399%
- 투자 특징: 하락장 회피

기타 암호화폐 시세 예측

종목	Accuracy	Precision	Recall	F1 Score
DOGE (도지코인)	0.8495	0.9057	0.8421	0.8727
Solana (솔라나코인)	0.8438	0.7368	1.0000	0.8485

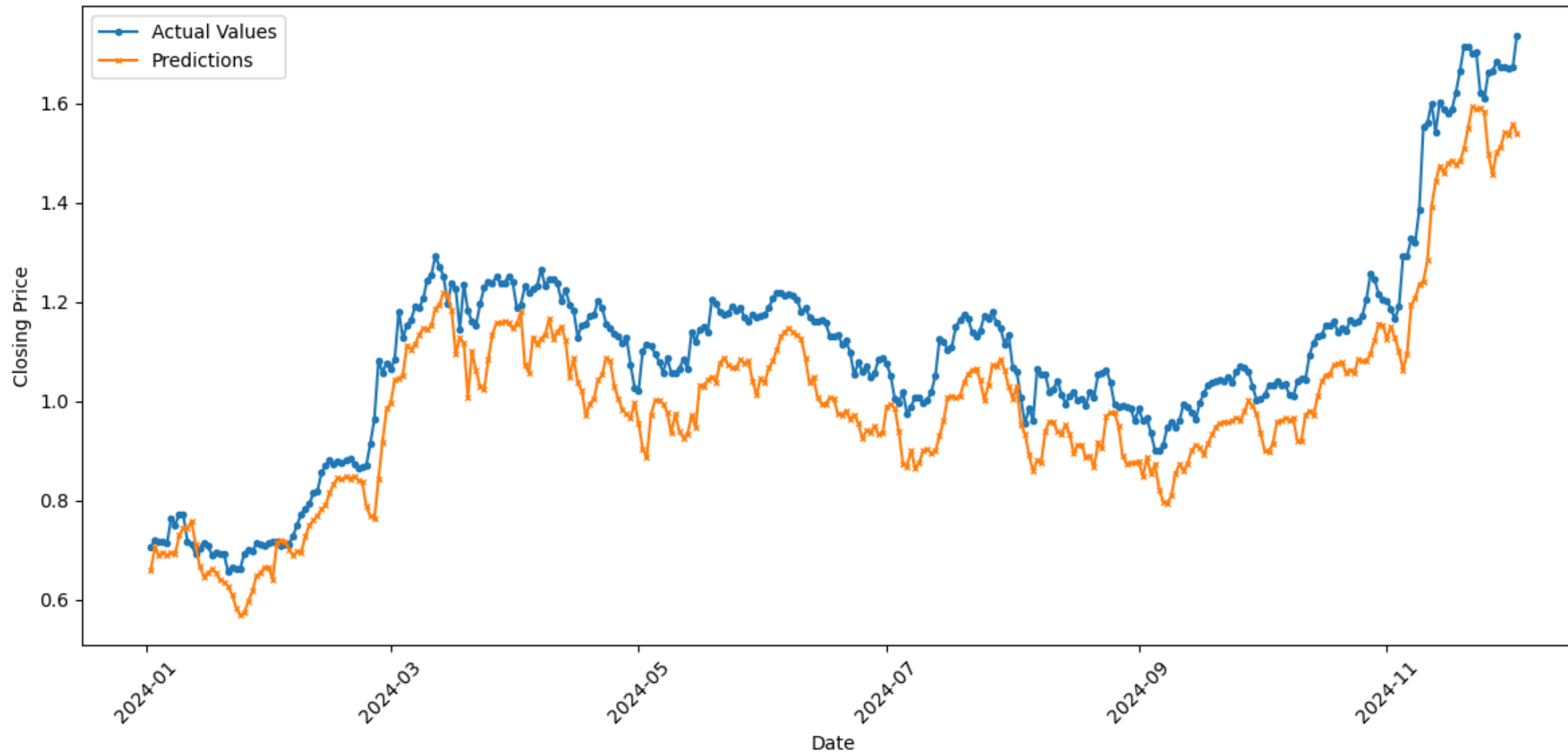


04

결론

결론

DNN Model: Actual vs. Predicted Closing Prices (2024)



암호화폐 가격 예측 모델 성능 분석

- 시계열 모델 대비 DNN이 더 우수한 성능
- 단순하지만 깊이 있는 학습 모델이 변동성 큰 암호화폐 예측에 적합
- 회귀모델 분석 결과, 종가 예측 시 전반적인 추세는 유사하나 실제 가격 크기의 큰 편차로 인해 투자 시 손실 위험이 높음

개선할 점

- 상승일지 하락일지만 판단할 뿐 얼마나 상승할지 하락할지는 알 수 없음.
- 만약 20%의 확률로 예측을 잘못하여 가격이 하락할 때 하락장이라면 손해가 막심 할 것.



05

소감 및 부록

소감

박훈석

비트코인 가격 예측 프로젝트를 진행하면서 이전 머신러닝 프로젝트의 아쉬웠던 점들을 보완할 수 있었습니다. 새로운 파생 변수를 설정하고 모델링하는 과정에서 도메인 지식을 쌓으며 투자의 방향성을 찾아갔고, 실제 코인 거래를 통해 이성적이고 객관적인 투자 전략의 중요성을 깨달았습니다.

이 현

머신러닝 프로젝트를 할 때는 부족한 시간으로 인해서 아쉬움이 많았지만, 딥러닝 프로젝트에서는 비교적 충분한 시간을 통해 우리 조가 원하는 결과물에 근접하여 성과가 나온 것 같습니다. 또한, 실제로 비트코인을 비롯한 다양한 코인 거래를 통해 이성적이고 객관적인 투자 전략 수립의 중요성을 깨달을 수 있었습니다.

정민관

비트코인 가격의 변동을 예측하는 모델을 학습시켜보면서 매매에 사용되는 기술 지표들도 공부할 수 있어서 도움이 많이 되었고 안전한 투자를 위해 모델을 더 개선하고 싶습니다.

신영섭

이전 머신러닝 프로젝트에서 아쉬웠던 점들을 보완하는 과정이 의미있었다고 생각합니다. 새로운 파생 변수를 설정하고 모델링 하는 작업이 기억에 남고, 도메인 지식을 공부하면서 투자의 방향성을 잡은 것 같습니다!

부 록

**본 정보는 투자 권유를 목적으로 하지 않으며,
투자 결정은 본인의 판단에 따라 신중히 이루어져야
합니다.**

기사 및 논문 출처

서지희 기자, “국내 가상자산 투자 일평균 거래대금 15조 육박…코스피·코스닥 합산 수준”, 이투데이, 2024.12.25, <https://www.etoday.co.kr/news/view/2432219>

조미현 기자, “금융 CEO 절반 “가상자산 투자 허용되면 매수할 것”, 한경, 2024.04.18 <https://www.hankyung.com/article/2024041811351>

김미희 기자, 어엿한 주류금융… 비트코인 ETF 운용자산 1000억弗 돌파, 파이낸셜 뉴스, <https://www.fnnews.com/news/202412301804111482>

신민경 기자, “서유석 금투협회장 “가상자산 ETF 추진”, 한경, 2025.01.02, <https://www.hankyung.com/article/2025010208416>

서유범, 황창하, Predicting Bitcoin Market Trend with Deep Learning Models, Quantitative Bio-Science, 2018.05.25

안유진, 오하영, On-Chain Data를 활용한 LSTM 기반 비트코인 가격 예측, JKIIICE, 2021.10

가즈아!!!!!!!!!!!!!!!!!!!!



끝.