


# 캡스톤디자인 중간보고서

제 목	국문	빅데이터 분석 기반 비트코인 가격 예측 모델		
	영문	Bitcoin price prediction model based on big data analysis		
진 행 상 황	중요마일스톤	1) 비트코인 가격 데이터 호출 : 업비트 api 등록 및 호출 후 csv 파일로 저장 2) 보조지표 계산 : ta-lib을 활용, 계산 소스코드 작성 3) 보조지표 데이터 전처리 : 시각화, 상관계수 등 전처리 위한 소스코드 작성 4) LSTM 모델 훈련 및 예측 : 모델 학습을 위한 소스코드 작성		
	진행상황	1) 비트코인 가격 데이터 호출 : api 등록 및 파일 저장 완료 2) 보조지표 계산 : 보조지표 계산 완료 3) 보조지표 데이터 전처리 : 시각화, 상관계수 완료, 기술 및 시계열 분석은 진행 예정 4) LSTM 모델 훈련 및 예측 : 전 단계 완료 후 진행 예정		
산출물	요구사항 정의서(별첨 1), 중간보고서(별첨 2)			
팀 구성원	학년	학 번	이 름	연락처(전화번호/이메일)
	4	20181590	김희섭	010-3338-6292/djlucifer@naver.com
	4	20181581	강현욱	010-8808-1182/20181581@edu.hanbat.ac.kr
컴퓨터공학과와 프로젝트 관리규정에 따라 다음과 같이 요구사항 정의서와 중간보고서를 제출합니다  <div style="display: flex; justify-content: space-between; align-items: flex-end;"> <div>2023 년 4 월 25 일</div> <div>             책임자 : 김희섭               지도교수 : 장수영 (인)           </div> </div>				

[별첨1]

프로젝트명 : 빅데이터 분석기반 비트코인 가격 예측 모델 제작

# 소프트웨어 요구사항 정의서

Version 1.0

개발 팀원 명(팀리더): 김희섭  
강현욱

대표 연락처:010-3338-6292

e-mail: djlucifer@naver.com

## 목차

1. 개요
2. 시스템 장비 구성요구사항
3. 기능 요구사항
4. 성능 요구사항
5. 인터페이스 요구사항
6. 데이터 요구사항
7. 테스트 요구사항
8. 보안 요구사항
9. 품질 요구사항
10. 제약 사항
11. 프로젝트 관리 요구사항

## 요구사항 정의서에 사용되는 양식 설명

요구사항 고유번호(ID): 제안요청서에 정의된 요구사항에 대해 계약, 사업수행, 사업완료 및 검수까지 변경, 삭제, 수정 여부에 대한 추적관리를 위해 고유의 번호를 부여하도록 한다.

요구사항 구분 및 ID부여 규칙

요구사항 구분		ID 부여 규칙
시스템 장비 구성 요구사항	Equipment Composition Requirement	ECR-000
기능 요구사항	System Function Requirement	SFR-000
성능 요구사항	Performance Requirement	PER-000
인터페이스 요구사항	System Interface Requirement	SIR-000
데이터 요구사항	Data Requirement	DAR-000
테스트 요구사항	Test Requirement	TER-000
보안 요구사항	Security Requirement	SER-000
품질요구사항	Quality Requirement	QUR-000
제약사항	Constraint Requirement	COR-000
프로젝트 관리 요구사항	Project Mgmt. Requirement	PMR-000
프로젝트 지원 요구사항	Project Support Requirement	PSR-000

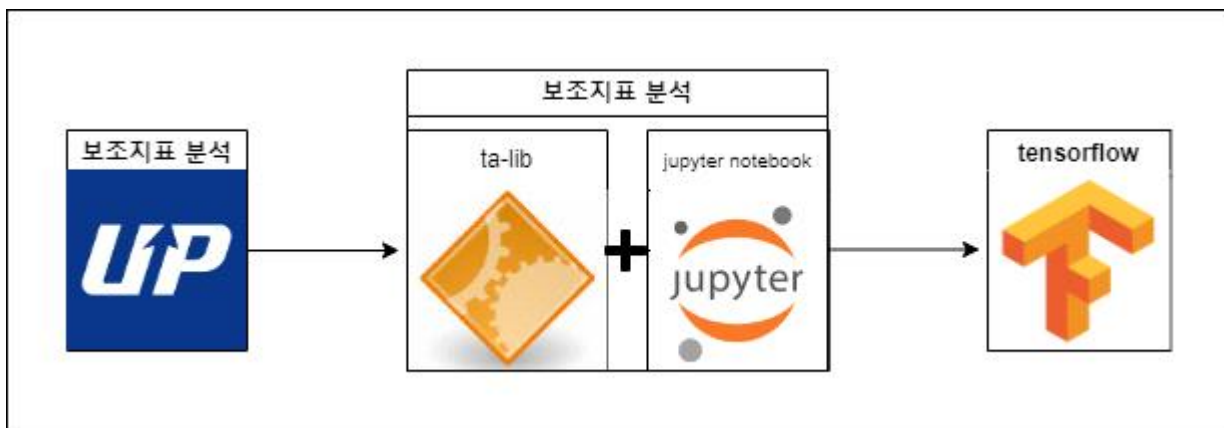
요구사항 세부내용 작성표 양식 및 항목설명

요구사항 고유번호		(설명) 요구사항 추적관리를 위해 독립적인 고유번호(ID) 부여
요구사항 명칭		(설명) 요구사항 명칭을 작성함
요구사항 분류		(설명) 요구사항 분류기준에 따른 분류를 기입
요구사항 상세 설명	정의	(설명) 요구사항 정의
	세부 내용	(설명) 요구사항 구체적인 세부 내용을 설명
산출정보		(설명) 해당기능을 통해 산출되는 결과물 혹은 정보를 표기
관련 요구사항		(설명) 정의된 요구사항과 관련된 요구사항에 대해 기술
요구사항 출처		(설명) 기능 도출내용에 대한 출처(source) 표기

## 1. 시스템 개요

목표 : 시각화, 상관관계수 등을 이용해 여러 보조지표들을 분석, 예측 모델 학습에 적합한 보조지표들을 추려낸 다음, LSTM 모델을 학습하여 비트코인 가격을 예측하는 모델 구현

업비트 api로부터 비트코인 가격 데이터 수집 -> ta-lib과 jupyter notebook을 이용해 지표 산출 및 분석 -> tensorflow를 통해 LSTM 모델 훈련



## 2. 시스템 장비 구성요구사항

- » 장비품목 : Jupyter notebook, Tensorflow
- » 장비 수량 : 단일
- » 장비 기능 : 데이터 분석, 훈련
- » 장비 성능 및 특징 : 데이터 분석에 사용되는 오픈 소스 웹 어플리케이션 및 머신러닝을 위한 오픈소스 플랫폼

### 3. 기능 요구사항

요구사항 고유 번호	SFR-001		
요구사항 명칭	비트코인 가격 데이터 호출		
요구사항 분류	기능	응락수준	필수
요구사항 상세 설명	정의	업비트 API를 이용해 비트코인 가격 데이터 수집	
	세부 내용	1. open, high, low, close, volume, value 값을 수집 2. 수집한 데이터를 csv 파일로 저장	

요구사항 고유 번호	SFR-002		
요구사항 명칭	보조지표 계산		
요구사항 분류	기능	응락수준	필수
요구사항 상세 설명	정의	수집한 가격 데이터를 이용해 보조지표 계산	
	세부 내용	ta-lib 패키지를 이용해 각 보조지표들을 계산	

요구사항 고유 번호	SFR-003		
요구사항 명칭	보조지표 데이터 전처리		
요구사항 분류	기능	응락수준	필수
요구사항 상세 설명	정의	보조지표 데이터 가공	
	세부 내용	1. 시각화 2. 실제 가격과의 상관계수 계산 3. 기술 및 시계열 분석	

요구사항 고유 번호	SFR-004		
요구사항 명칭	LSTM 모델 훈련 및 예측		
요구사항 분류	기능	응락수준	필수
요구사항 상세 설명	정의	시분할 모델 중 LSTM 모델을 이용해 훈련	
	세부 내용	1. 학습용 데이터로 선정된 지표들을 여러 방식으로 조합해 보며 학습 2. 파라미터 최적화	

#### 4. 성능 요구사항

요구사항 고유 번호	PER-001		
요구사항 명칭	예측 정확도 및 소요 시간		
요구사항 분류	성능	응락수준	필수
요구사항 상세 설명	정의	예측 모델의 정확도 및 소요 시간	
	세부 내용	1. 학습 시 학습시킬 보조지표들의 조합을 통해 최적의 정확도 보이도록 함 2. 예측 범위에 따른 소요시간을 조절(분봉 예측 시 예측 소요 시간이 이를 넘어서는 안됨)	

## 5. 인터페이스 요구사항

요구사항 고유 번호	SIR-001		
요구사항 명칭	예측 결과 시각화		
요구사항 분류	인터페이스	응락수준	필수
요구사항 상세 설명	정의	모델의 예측을 확인하기 위한 결과의 시각화	
	세부 내용	모델의 예측값과 실제 비트코인 가격 디스플레이	

## 6. 데이터 요구사항

요구사항 고유 번호	DAR-001		
요구사항 명칭	데이터 저장		
요구사항 분류	데이터	응락수준	필수
요구사항 상세 설명	정의	수집한 비트코인 가격 데이터 및 학습용 보조지표 데이터 저장	
	세부 내용	1. 보조지표 별 각각 따로 데이터를 저장 2. 각 데이터는 csv 파일로 저장	



7. 테스트 요구사항

요구사항 고유 번호	TER-001		
요구사항 명칭	예측 모델 테스트		
요구사항 분류	테스트	응락수준	필수
요구사항 상세 설명	정의	예측값의 정확도 테스트	
	세부 내용	1. 실제 비트코인 가격 시각화 2. 예측값 시각화 3. 두 값을 비교하여 파라미터 값을 조정하여 최적화	

8. 보안 요구사항

요구사항 고유 번호	SER-001		
요구사항 명칭	API 발급 키 보안		
요구사항 분류	보안	응락수준	필수
요구사항 상세 설명	정의	API 발급을 위한 개인키 보안	
	세부 내용	예측 모델 학습과 추후 성능 개선에 쓰일 업비트 api 공개키 및 암호키 비공개	

## 9. 품질 요구사항

요구사항 고유 번호	QUR-001		
요구사항 명칭	예측 모델의 실용성		
요구사항 분류	품질	응락수준	필수
요구사항 상세 설명	정의	실제 투자에서의 적합성	
	세부 내용	1. 예측 모델을 통한 비트코인 모의투자를 통해 수익률 기록 2. 비트코인의 변동성을 고려, 최소한의 손실률을 기록	

## 10. 제약 사항

요구사항 고유 번호	COR-001		
요구사항 명칭	개발에서의 제약사항		
요구사항 분류	제약사항	응락수준	필수
요구사항 상세 설명	정의	개발 과정에서의 제약사항	
	세부 내용	1. 여러 보조지표들을 학습 데이터로 사용하기 위해 각 지표 값을 통일할 필요성 2. jupyter notebook, tensorflow 이용을 위한 python 사용 3. 지속해서 변동하는 비트코인 가격으로 인한 지속적인 데이터 갱신 필요	

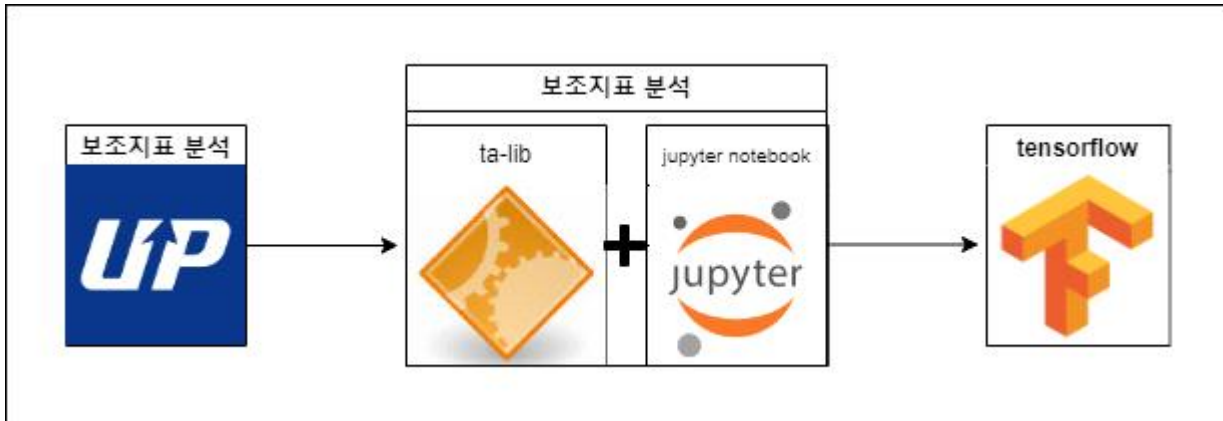
## 11. 프로젝트 관리 요구사항

요구사항 고유 번호	PMR-001		
요구사항 명칭	품질관리		
요구사항 분류	프로젝트 관리	응락수준	필수
요구사항 상세 설명	정의	프로젝트 관리	
	세부 내용	1. 우선적으로 예측 모델을 만들기 위한 데이터셋 구축 2. 데이터셋을 여러 방식으로 조합해 보며 최적의 예측 모델 결정 3. 추후 예측 정확도를 높이기 위한 추가적인 데이터셋을 구축하거나 자동 매수, 매도 시스템을 추가	

## [별첨2]

# 중간보고서

1. 요구사항 정의서에 명시된 기능에 대하여 현재까지 분석, 설계, 구현(소스코드 작성) 및 테스트한 내용을 기술하시오.



- 1) 업비트 api로부터 비트코인 가격 데이터 수집

```
1 df = up.get_ohlcv(ticker="KRW-BTC", interval="day", count=200)
2 print(df)
```

	open	high	low	close	#
2022-09-22 09:00:00	26523000.0	27699000.0	26478000.0	27572000.0	
2022-09-23 09:00:00	27571000.0	27788000.0	26709000.0	27594000.0	
2022-09-24 09:00:00	27591000.0	27611000.0	27101000.0	27270000.0	
2022-09-25 09:00:00	27270000.0	27499000.0	26910000.0	27051000.0	
2022-09-26 09:00:00	27051000.0	27749000.0	27000000.0	27597000.0	
...	...	...	...	...	...
2023-04-05 09:00:00	37207000.0	37804000.0	36948000.0	37139000.0	
2023-04-06 09:00:00	37139000.0	37211000.0	36800000.0	37058000.0	
2023-04-07 09:00:00	37058000.0	37200000.0	36890000.0	36993000.0	
2023-04-08 09:00:00	36993000.0	37249000.0	36930000.0	37080000.0	
2023-04-09 09:00:00	37064000.0	37181000.0	37032000.0	37139000.0	
...	...	...	...	...	...
2022-09-22 09:00:00	5953.989094	1.614473e+11			
2022-09-23 09:00:00	7237.642087	1.975800e+11			
2022-09-24 09:00:00	2902.696438	7.949777e+10			
2022-09-25 09:00:00	2488.111300	6.783023e+10			
2022-09-26 09:00:00	4639.446161	1.270504e+11			
...	...	...	...	...	...
2023-04-05 09:00:00	4804.317546	1.798466e+11			
2023-04-06 09:00:00	3265.195674	1.208360e+11			
2023-04-07 09:00:00	2191.385527	8.117821e+10			
2023-04-08 09:00:00	3044.987229	1.130097e+11			
2023-04-09 09:00:00	369.936960	1.373656e+10			

```
[200 rows x 6 columns]
```

```
[ ] 1 df.to_csv('bitcoin_data.csv')
```

```
[ ] 1 df = pd.read_csv('bitcoin_data.csv')
```

- 2) 각 보조지표 계산 (상위 4개)

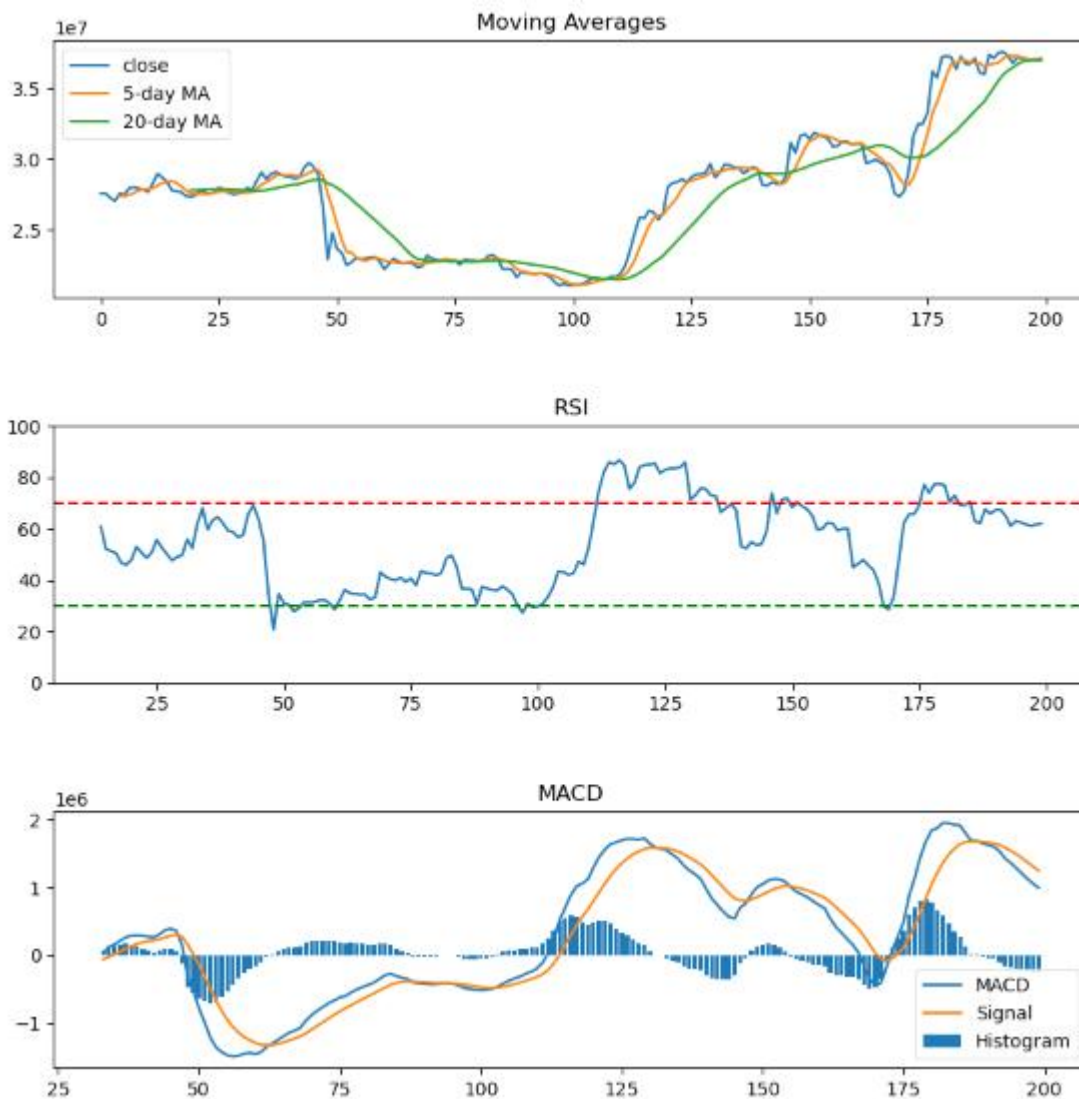
```
1 # 이동평균 계산
2 ma_5 = ta.SMA(df['close'], timeperiod=5) # 5일 이동평균
3 ma_20 = ta.SMA(df['close'], timeperiod=20) # 20일 이동평균
4
5 # RSI 계산
6 rsi = ta.RSI(df['close'], timeperiod=14)
7
8 # MACD 계산
9 macd, signal, hist = ta.MACD(df['close'], fastperiod=12, slowperiod=26, signalperiod=9)
10
11 # 볼린저 밴드 계산
12 upper, middle, lower = ta.BBANDS(df['close'], timeperiod=20, nbdevup=2, nbdevdn=2)
```

### 3) 각 지표 시각화 (상위 3개)

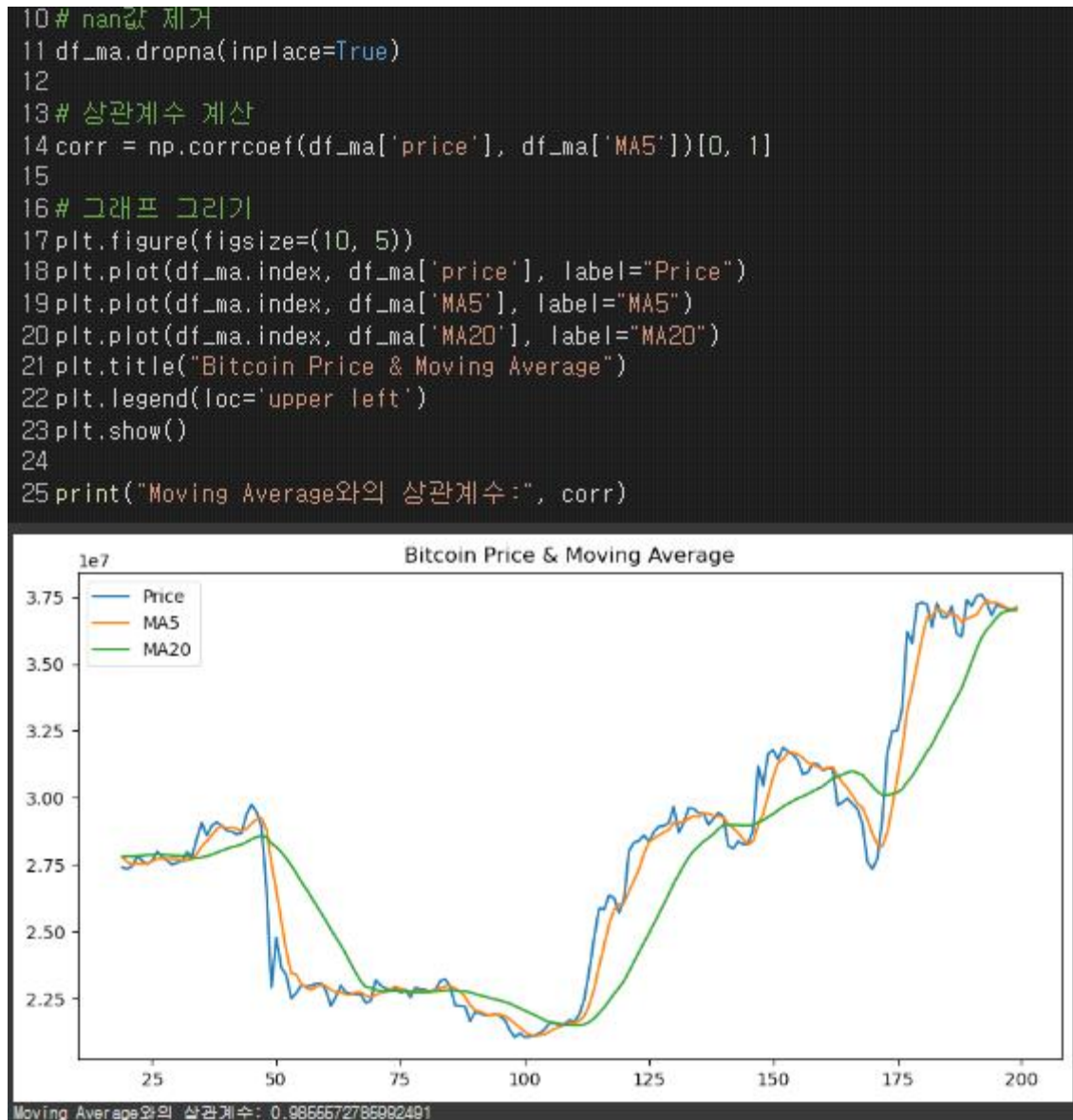
```

1 # 그래프 그리기
2 fig, axes = plt.subplots(nrows=3, ncols=1, figsize=(10, 10))
3 plt.subplots_adjust(hspace=0.5)
4
5 # 이동평균 그래프
6 axes[0].plot(df['close'], label='close')
7 axes[0].plot(ma_5, label='5-day MA')
8 axes[0].plot(ma_20, label='20-day MA')
9 axes[0].set_title('Moving Averages')
10 axes[0].legend()
11
12 # RSI 그래프
13 axes[1].plot(rsi)
14 axes[1].axhline(y=30, color='g', linestyle='--')
15 axes[1].axhline(y=70, color='r', linestyle='--')
16 axes[1].set_title('RSI')
17 axes[1].set_ylim([0, 100])
18
19 # MACD 그래프
20 axes[2].plot(macd, label='MACD')
21 axes[2].plot(signal, label='Signal')
22 axes[2].bar(df.index, hist, label='Histogram')
23 axes[2].set_title('MACD')
24 axes[2].legend()
25

```



4) 실제 비트코인 가격 간의 상관관계수 계산 (실제 가격과의 그래프 비교, 이동평균선 기준)



2. 프로젝트 수행을 위해 적용된 추진전략, 수행 방법의 결과를 작성하고, 만일 적용과정에서 문제점이 도출되었다면 그 문제를 분석하고 해결방안을 기술하시오.

1) 몇몇 보조지표들 중 ta-lib 호출 함수가 python 버전에 따라 사용이 불가능한 상황이 발생

-> 처음에는 사용불가능한 보조지표들을 제외한 뒤 나머지 지표들로만 분석을 진행, 분석을 진행하면서 많은 보조지표 중 비트코인과 인과관계가 높은 지표들이 많지 않다고 생각되어 그대로 진행함

2) 시각화에 있어서 각 지표별로 나타내는 값이 달라 실제 비트코인 가격 그래프와의 비교에 어려움 발생

-> 값을 맞출 수 있는 그래프는 맞추어 진행, 나머지는 따로 그래프를 만들어 비교

3) 상관관계수 계산 시 NaN 값으로 인한 오류

-> dropna 함수를 사용하여 dataframe에 있는 NaN 값을 제거(이 작업을 안할 경우 오류)

# 캡스톤 디자인 | 중간보고서 채점표

평가도구	평 가 항 목	평 가 점 수				
		1	2	3	4	5
중간 보고서 및 실행 결과	1. 요구사항 정의서(기능, 성능, 인터페이스 등)가 구체적으로 작성되었는가?					
	2. 요구분석, 설계 산출물(모델, 프로토타입 등)의 내용이 충실한가?					
	3. 설계 및 구현 문제를 위해 적용한 이론, 문제해결 방법이 제시되었으며 그 적용이 적합한가?					
	4. 구현된 소프트웨어(또는 이와 동등한 하드웨어 시스템)가 버그없이 실행되었는가?					
	5. 구현된 소프트웨어(또는 이와 동등한 하드웨어 시스템)의 성능 요구사항은 충족되었는가?					
도구활용	6. 설계 및 구현을 위해 도구가 적절히 활용되었는가?					
	7. 도구의 활용수준(능숙도)은 프로젝트 수행에 적합한가?					
팀원의 업무 및 역할	8. 팀원의 업무분담에 따른 역할 및 협력이 충실히 이루어졌는가? (평가자에 의한 질의)					
	9. 프로젝트 중간 진척상황에 대해 팀원이 충분히 인지하고 있는가?(평가자에 의한 질의)					
합계						
*검토 의견(최종완료 때까지 보완해야할 점에 대해 작성 요망)						
심사위원(소속):		(이름)			(인)	