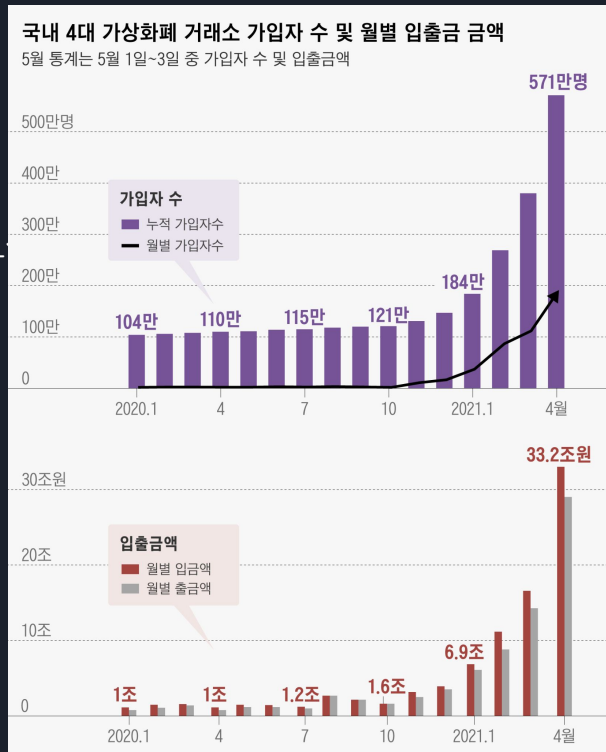


# Project 4 - Bitcoin 분석 딥러닝 모델

AI10기 임규현

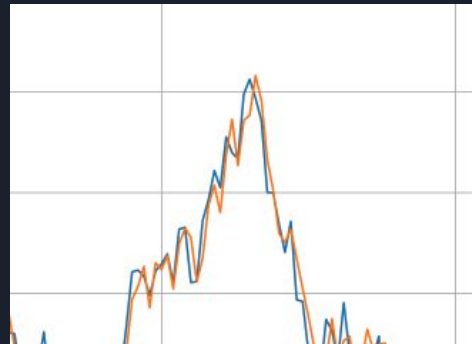
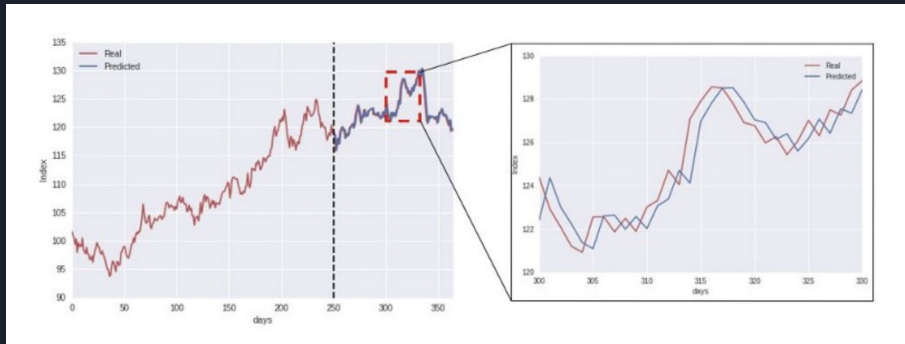
# Background

- 비트코인은 암호화폐의 나스닥 / 코스피 같은 존재
  - 암호화폐 시장의 약 40% 를 차지
- 늘어나는 기관 및 대중의 참여
- Q. 딥러닝 모델이 비트코인 예측에 적용될 수 있을까
- 매매에 대한 기준? 이 있는가?
  - 일반인들은 기분, 뉴스, 지인, 차트 등을 사용



# Problem Statement - 1

## 1. 일반적으로 온라인에서 접할 수 있는 관련 딥러닝 모델들의 문제



1.1. 얼핏 보면 쉽게 돈을 벌 수 있을 것 같다. 하지만?

1.2. Overfitting? 왜? -> **Loss 값의 최적화** 때문

1.2.1. 어떤 최적화? - 어제의 값이 내일의 값과 같다

# Problem Statement - 2

## 1. 대안?

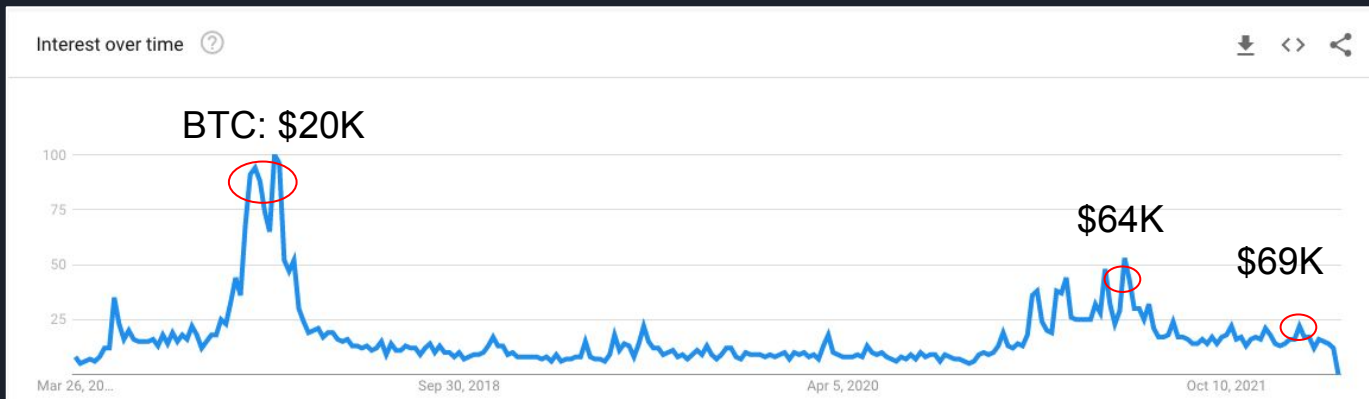
1.1. Non-stationary -> Stationary Data로의 전환 -> 모델링

1.2. 내일의 가격이 **상방**일지 **하방**일지 예측 (이중분류)

1.3. Feature?

1.3.1. Google Trend, 뉴스, 차트, 유동성, Twitter Sentiment Analysis

(+, -)



# Problem Statement - 2

## 1. 대안?

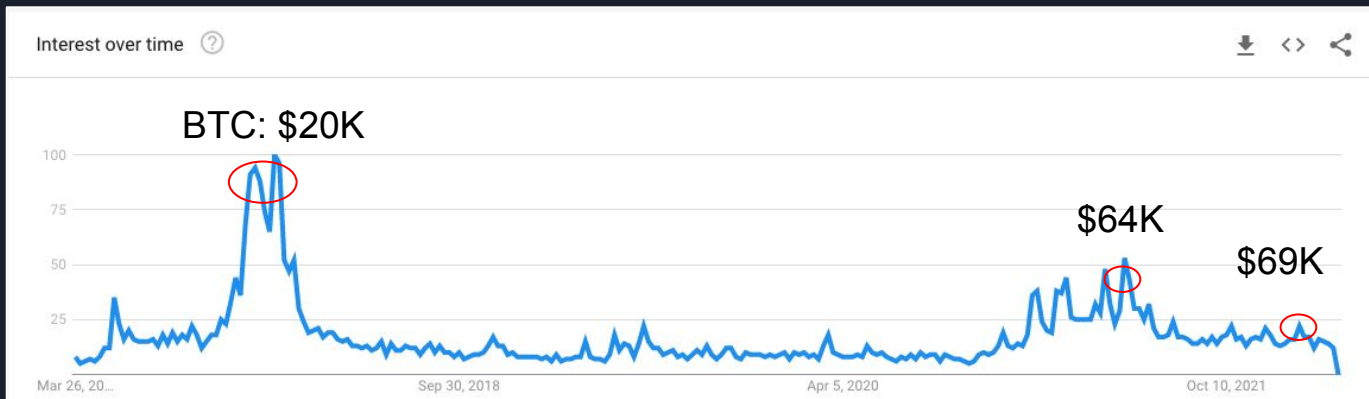
1.1. Non-stationary -> Stationary Data로의 전환 -> 모델링

1.2. 내일의 가격이 상방일지 하방일지 예측 (이중분류)

## 1.3. Feature?

1.3.1. Google Trend, 뉴스, 차트, 유동성, Twitter Sentiment Analysis

(+, -)





# 데이터 파이프라인

데이터  
수집

데이터  
전처리

모델링

모델 분석  
및 검증

모델 배포



# 데이터 파이프라인

데이터  
수집

데이터  
전처리

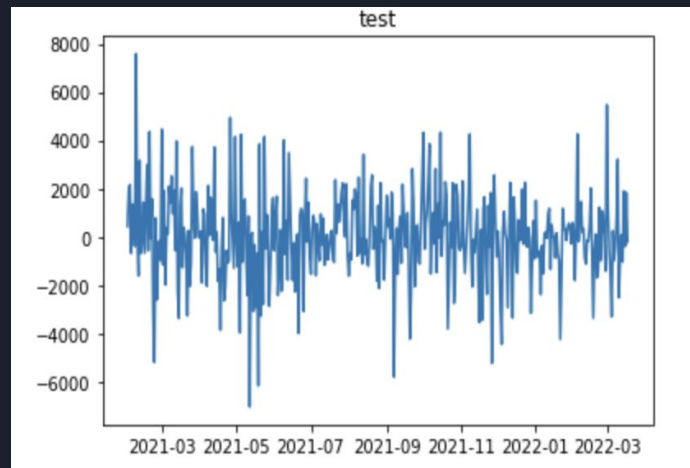
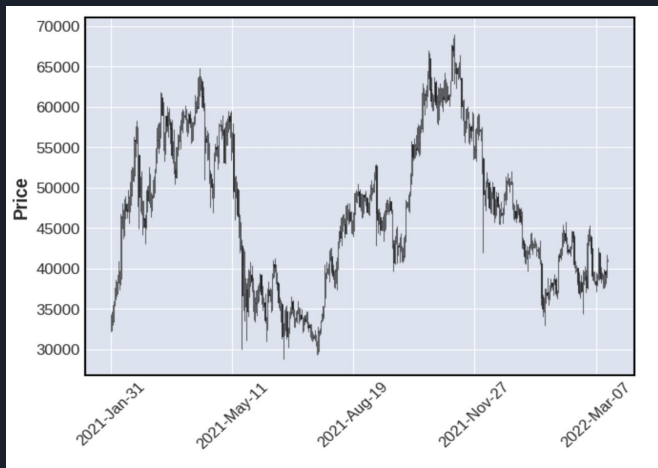
모델링

모델 분석  
및 검증

모델 배포

# Data Input & Processing - 1

- Alpha Vantage 의 API를 사용
  - 일별 비트코인 가격 데이터를 받아옴 (시가,저가,종가,거래량)
- 기간 설정 - 2021년 1월
- 종가 예측
  - Non-Stationary -> Stationary 로 변환 (오늘 가격 - 전일 가격); Dickey-Fuller?





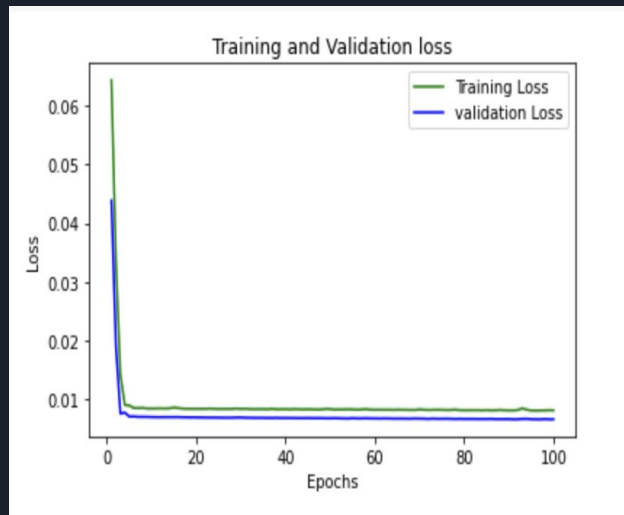
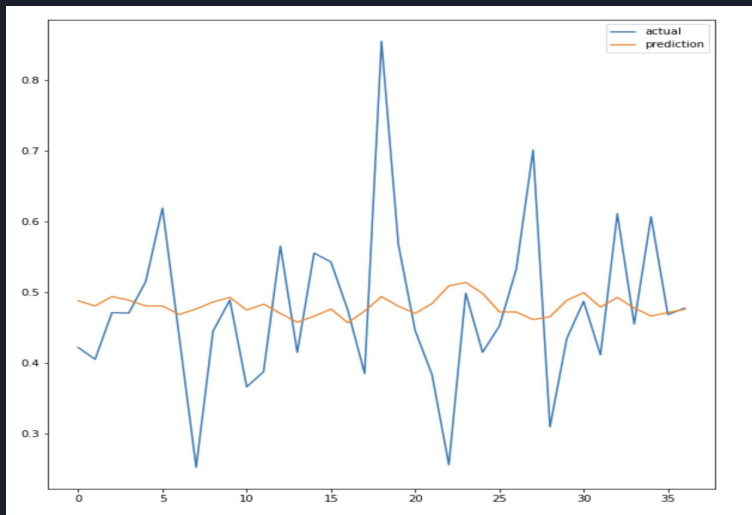


# Data Input & Processing - 2

- 이진분류 모델
  - 종가 ('Close') 전처리
    - 종가가 +로 마감 -> 1, -로 마감 -> 0
- 모델의 성능을 높이기 위해 정규화 (0~1)
  - 며칠간의 데이터를 보고 그 다음 것을 예측할지 window를 설정

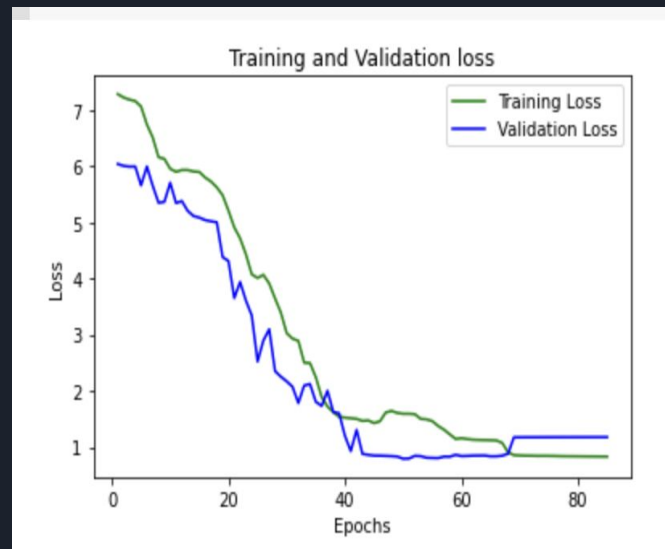
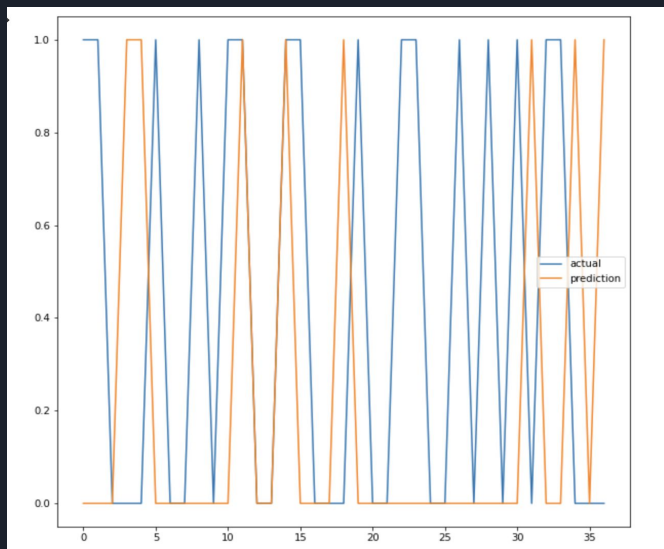
# Modeling & Evaluation -1

- 기본적으로 LSTM을 사용
  - 이전 데이터 정보 일부를 다음 분석에 넘겨줌
    - 시계열 데이터 (Sequential Data)에 적합
- 첫 번째 모델 - Stationary Data에 대한 LSTM

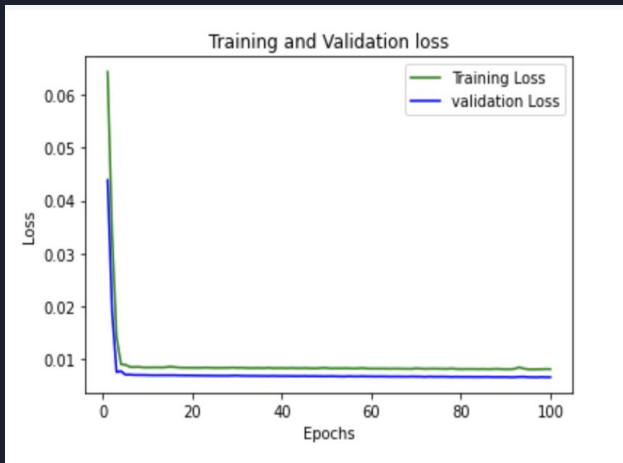


# Modeling & Evaluation -2

- LSTM
  - 이진분류

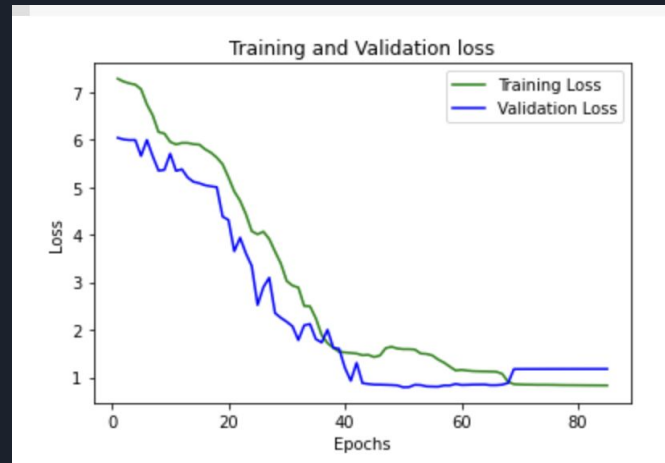


# Evaluation



```
model.evaluate(test_data, verbose=1)
```

```
2/2 [=====] - 0s 35ms/step - loss: 0.0067 - mse: 0.0134  
[0.006689791567623615, 0.01337958313524723]
```



```
2/2 [=====] - 0s 35ms/step - loss: 0.7913 - mse: 0.2917  
[0.7912996411323547, 0.291693776845932]
```



# Limitations

- 서비스 개발?
- 오직 가격을 기반으로 한 모델링
- 딥러닝 모델 디자인
- 모델 검증
  - 하이퍼 파라미터 튜닝

# Improvements

- Sentiment Analysis
- Technical Analysis (차트 패턴에 기반한)
- Modeling - ARIMA, ETS 와 같은 다른 모델들을 사용