第四組
推特情緒分析與
股價報酬的影響

312707020 林震洋 財金所碩一
110705018 朱建丞 資財系大三
109705059 蔡思妤 資財系大四

TABLE OF CONTENTS

1

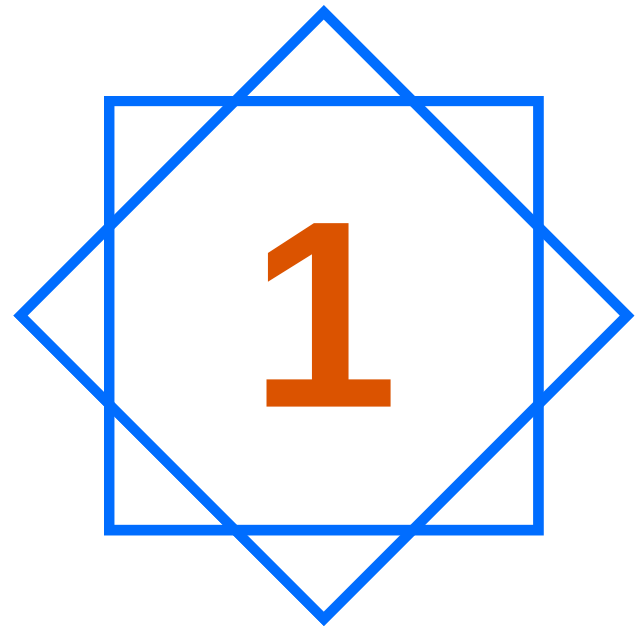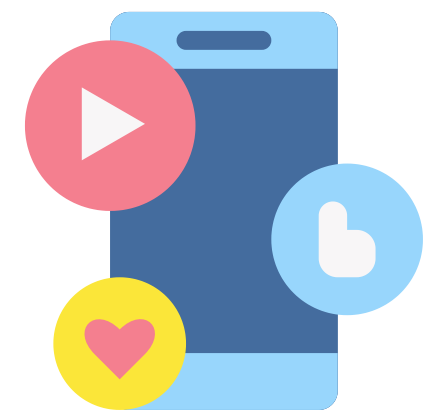# INTRODUCTION

INTRODUCTION

- 透過情緒分析套件**VADER**（Valence Aware Dictionary for Sentiment Reasoning）model評價Twitter上的推文，並使用**GANs**(Generative Adversarial Networks)模型預測股價，最後和傳統技術分析比較。

2

SENTIMENT ANALYSIS

WHY IS IT SO DIFFICULT TO CONDUCT A SENTIMENT ANALYSIS ?

1

Sarcasm

指說的話與你想說的截然相反。

2

Decoding

句子的主詞和受詞是什麼，動詞或形容詞與誰有關？

3

Tweets

火星文、拼字錯誤、缺少標點符號、奇怪的文法結構。

PRE-PROCESS AND DATA CLEANING

1

Tokenization

2

Removing stop words

3

Normalizing words

4

Text Vectorization

VADER
(VALENCE AWARE DICTIONARY
AND SENTIMENT REASONER)

- It is a lexical database and rule-based sentiment analysis tool that is **optimized for social media sentiments.**

- It not only **exhibits the Positivity and Negativity scores**, but also the degree to which a sentiment is positive or negative

|	Date	Tweet	Stock Name	sentiment_score	Negative	Neutral	Positive
0	2022/12/31	Huh. So Intel discontinued Optane, which means...	INTV	0.0	0.0	1.0	0.0
1	2022/12/31	A ?§µl was a pilot on these missions. I respec...	INTV	0.2617	0.0	0.945	0.055
2	2022/12/31	Christmas gift from the hubs. https://t.co/CAW...	INTV	0.4404	0.0	0.633	0.367
3	2022/12/31	German intelligence didn? t actually notice th...	INTV	0.6705	0.057	0.746	0.198
4	2022/12/31	? ¨INTEL? ¨\n\nOpTic Texas will have at least ...	INTV	0.0	0.0	1.0	0.0

3

GAN MODEL

GANS (GENERATIVE ADVERSARIAL NETWORKS)
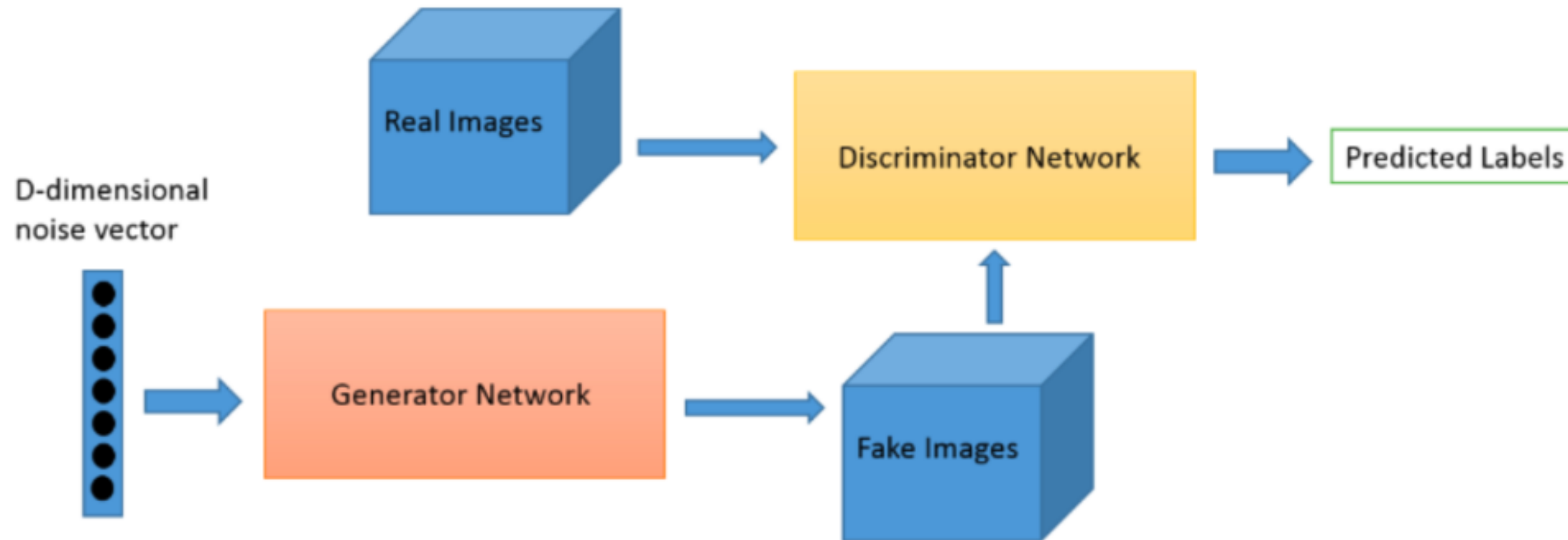
- Generator 會不斷想辦法"騙過" Discriminator。而Discriminator 會努力的去判別。所以雙方會不斷的進步成長。
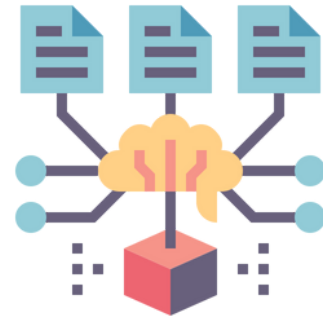
GANS
(GENERATIVE ADVERSARIAL NETWORKS)
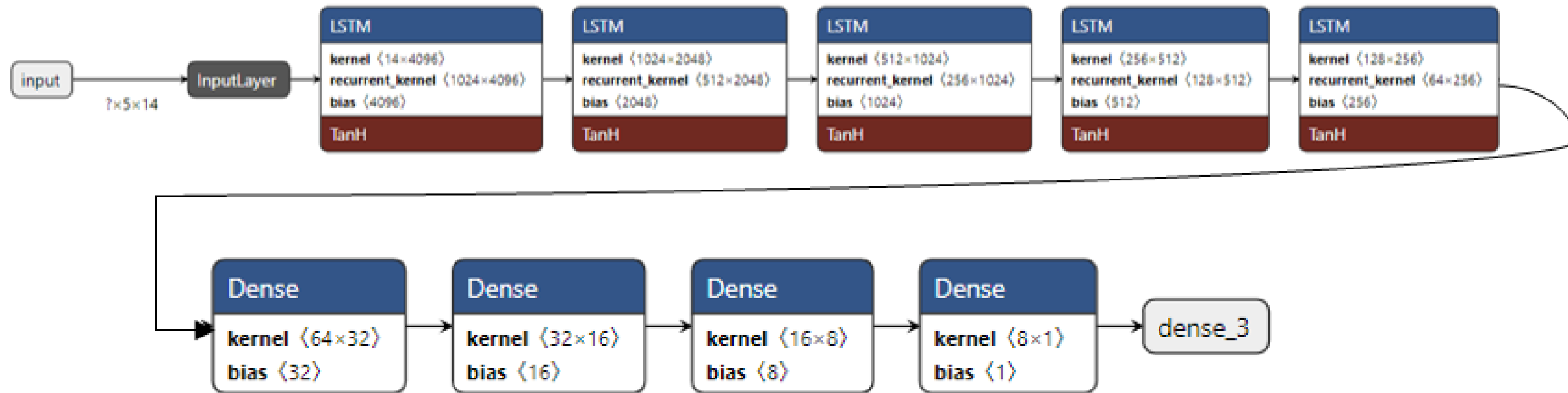
- 我們建立了一個 GAN 模型架構，其中generator 有 5 個 LSTM 區塊，discriminator有 5 個convolutional 和 3 個帶有 sigmoid 活化函數的dense layers。

4

RESULTS

訓練模型

- 使用資料 : 2018年Advanced Micro Device(AMD)的推特文章
- 訓練模型 : 使用GAN將以上特徵輸入進行訓練
- 預測結果 : 將訓練好的模型對2019年AMD的股價進行預測

Open	High	Low	Close	Adj Close	Volume	sentiment _score	MA7	MA20	MACD	20SD	upper_ band	lower_band	EMA	logmomentum

模擬股價

Prediction on test data for AMD
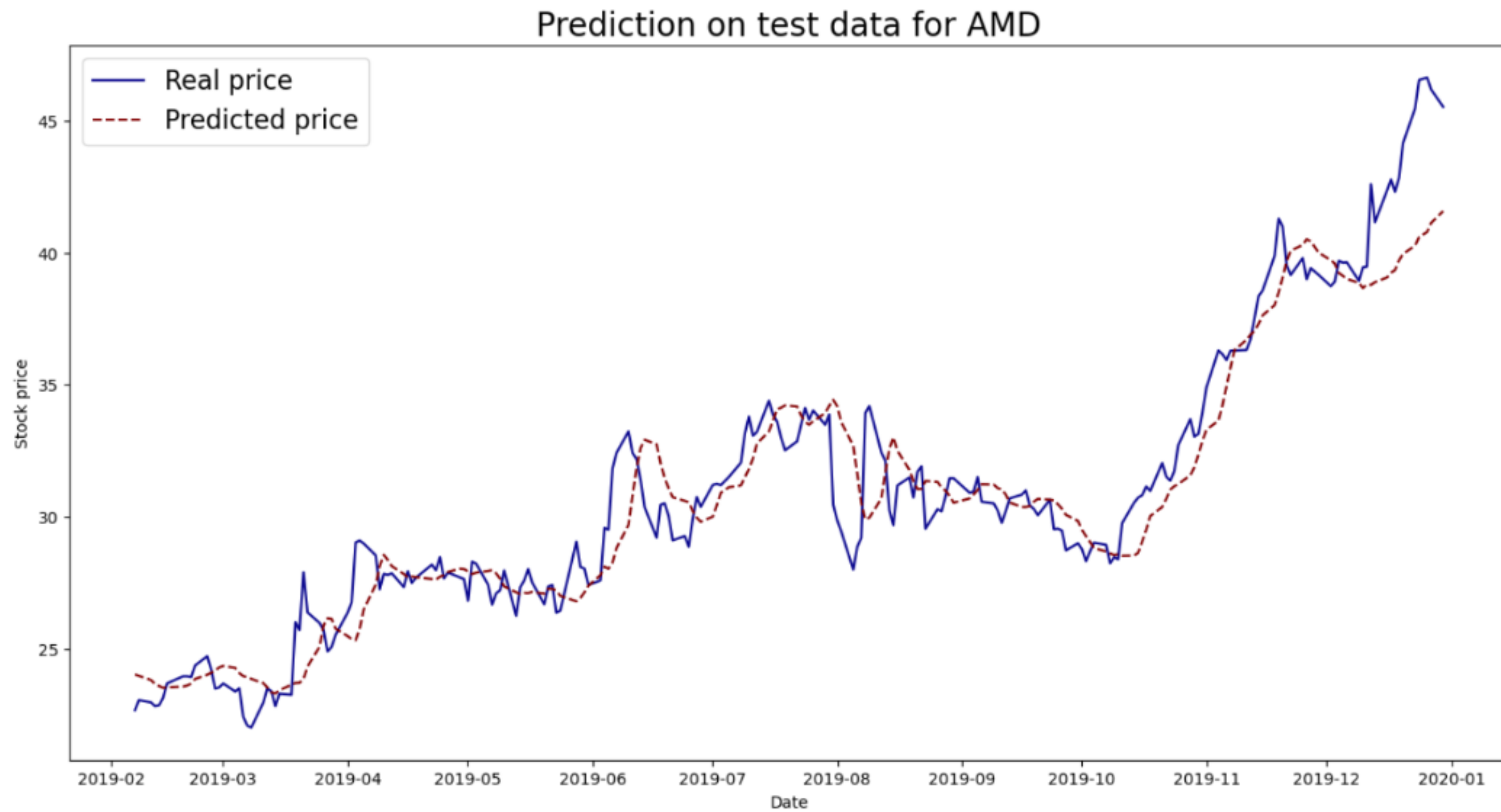
建立投資策略

- **做多：**

1. long condition : 當預測的股價大於前一天的收盤價，以開盤價進行買入
2. 停利 : 預測股價小於前一天收盤價
3. 停損 : 買入訊號的五天k棒最低點

- **做空：**

1. short condition : 當預測的股價小於前一天的收盤價，以開盤價進行賣出
2. 停利 : 預測股價大於前一天收盤價
3. 停損 : 買入訊號的五天k棒最高點

交易策略目標

- 夏普比率（Sharpe Ratio）：

$$Sharpe\ Ratio = \frac{R_p - R_f}{\sigma_p}$$

- 卡瑪比率（Calmar Ratio）：

$$Calmar\ Ratio = \frac{\underline{R} - Rf}{Max\ Drawdown}$$

- 總回報（Total Return）：

是指在某個時間段內投資組合或資產的總體回報，通常包括資本增值和收益（如股息、利息等）

SELL   BUY

RESULTS

```
Start                                                    0
End                                                    225
Period                              226 days 00:00:00
Start Value                                          100.0
End Value                                        235.302783
Total Return [%]                                 135.302783
Benchmark Return [%]                             100.695954
Max Gross Exposure [%]                               100.0
Total Fees Paid                                        0.0
Max Drawdown [%]                                  16.614402
Max Drawdown Duration               53 days 00:00:00
Total Trades                                            82
Total Closed Trades                                     81
Total Open Trades                                        1
Open Trade PnL                                     4.180767
Win Rate [%]                                      61.728395
Best Trade [%]                                    15.904366
Worst Trade [%]                                   -7.344262
Avg Winning Trade [%]                             3.472369
Avg Losing Trade [%]                              -2.674994
Avg Winning Trade Duration          2 days 11:31:12
Avg Losing Trade Duration   2 days 06:11:36.774193548
Profit Factor                                     1.769748
Expectancy                                         1.61879
Sharpe Ratio                                      3.099778
Calmar Ratio                                     17.953411
Omega Ratio                                       1.580226
Sortino Ratio                                     5.397807
```
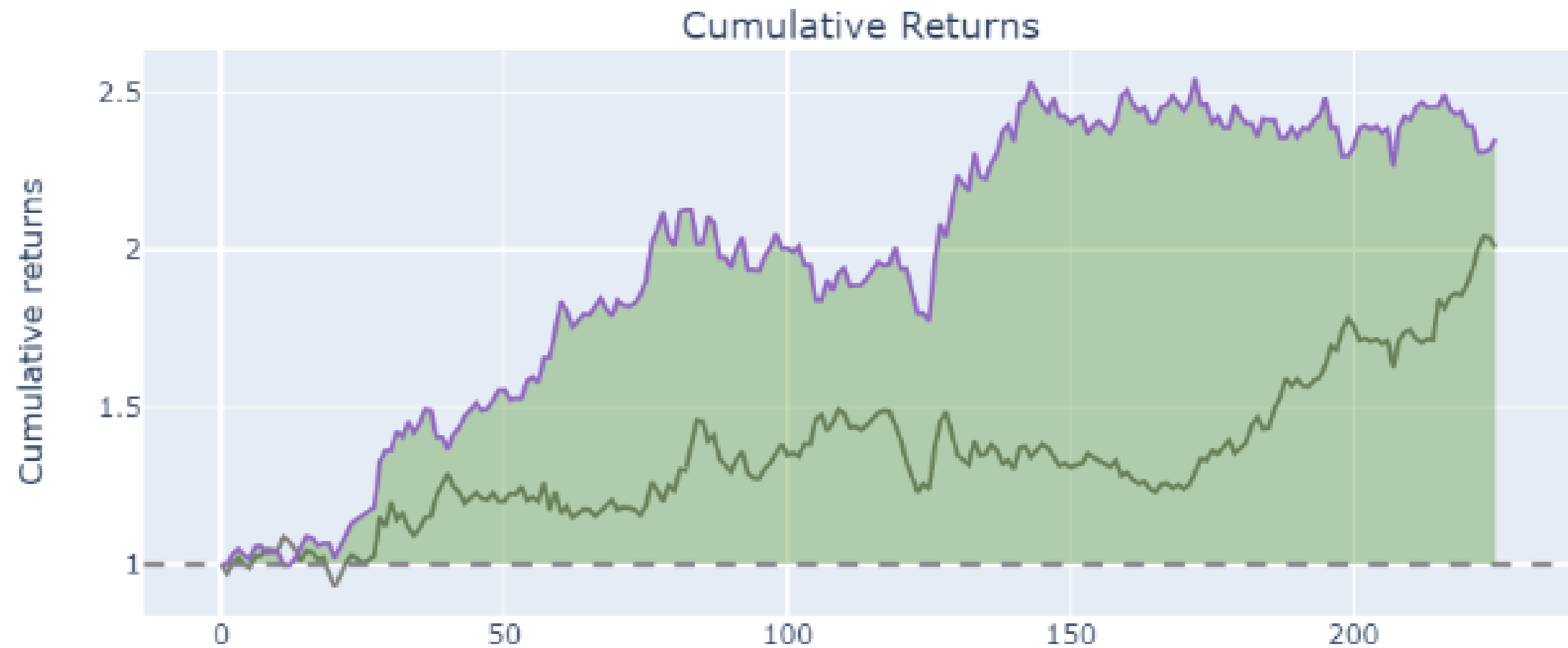
RESULTS

RESULTS

Trade PnL

RESULTS

- **Cumulative Return**

Cumulative Returns

THANK YOU