Stock movement as a Markov Process using Machine Learning predictor

Business Lab for Financial Engineering

Nguyen Minh Duc

20202026

Contents

- Trading Strategy
- Stock Movement as Markov Process
- Experiment

Trading Strategy

Just buy low and sell high, simple!

How?

Trading Strategy

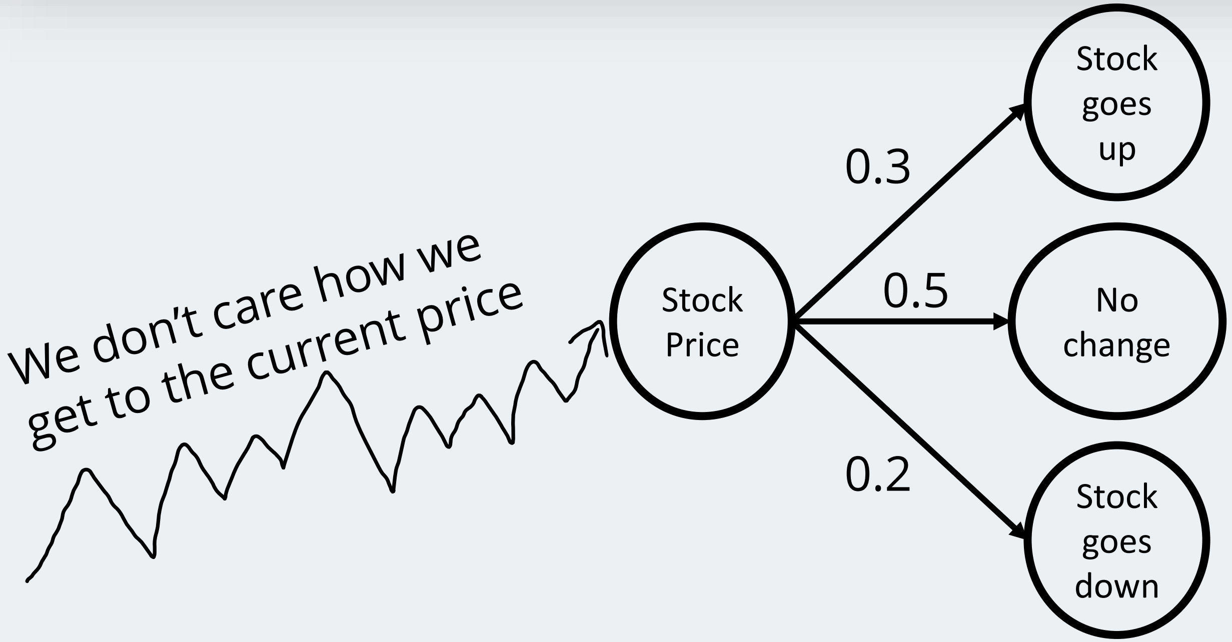
Just buy low and sell high, simple!

Then, how to predict the future?

Stock Movement as a Markov Process

1 What is a Markov process?

Given the historical states of a random system $X_1, X_2, \ldots, X_t$, the probability of moving to the next state depends only the current state, i.e., $P(X_{t+1} = x | X_1, X_2, \ldots, X_t) = P(X_{t+1} = x | X_t)$



Stock Movement as a Markov Process

2 Stock Movement as a Markov process

We consider the daily return of the stock: $\mathrm{r_t} = \frac{P_t}{P_{t-1}} - 1.$

Classify each $\mathrm{r_t}$ as

- High Increase (HI): $\mathbf{r_t} > \boldsymbol{Q_{inc}(75\%)}$
- Moderate Increase (MI): $\boldsymbol{Q_{inc}(50\%)} \leq \mathbf{r_t} < \boldsymbol{Q_{inc}(75\%)}$
- Slight Increase (SI): $\boldsymbol{Q_{inc}(25\%)} \leq \mathbf{r_t} < \boldsymbol{Q_{inc}(50\%)}$
- Neutral (Ne): $\boldsymbol{Q_{dec}(25\%)} \leq \mathbf{r_t} < \boldsymbol{Q_{inc}(25\%)}$
- Slight Decrease (SD): $\boldsymbol{Q_{dec}(25\%)} \leq \boldsymbol{r_t} < \boldsymbol{Q_{dec}(25\%)}$
- Moderate Decrease (MD): $\boldsymbol{Q_{dec}(25\%)} \leq \boldsymbol{r_t} < \boldsymbol{Q_{dec}(25\%)}$
- High Decrease (HD): $\boldsymbol{r_t} < \mathbf{Q_{dec}(25\%)}$

Stock Movement as a Markov Process

2 Stock Movement as a Markov process

We define a hyperparameter: *lookback* – The number of prior prices to consider
forward – The number of future days to predict

The sequence of prices can be encoded as a tuple based on the daily return
Example: (HI, SD Ne, HD, MI)

Now, we can use the historical prices to estimate the probability distribution of the next price state

$$P\left(X_{t+1} = \left(x_{t+1}, x_{t+2}, \dots, x_{t+forward}\right) \middle| X_t = \left(x_{t-lookback+1}, x_{t-lookback+2}, \dots, x_t\right)\right)$$

Stock Movement as a Markov Process

2 Stock Movement as a Markov process

How can we approximate this probability with the sample stock data?

$$P\left(X_{t+1} = \left(x_{t+1}, x_{t+2}, \ldots, x_{t+forward}\right) \middle| X_t = (x_{t-lookback+1}, x_{t-lookback+2}, \ldots, x_t)\right)$$

Previous approach was to use a table to count the frequency of every possible combination of price movements

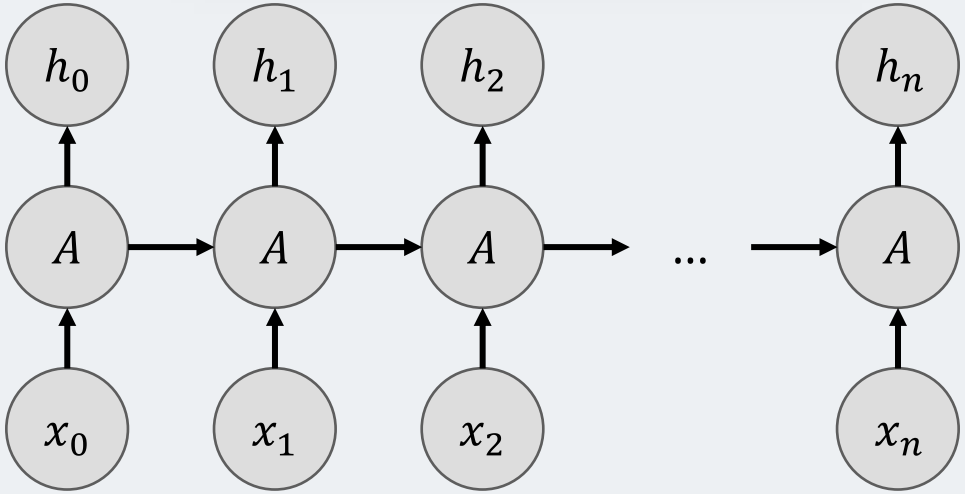
Memory complexity: $\boldsymbol{O}(\boldsymbol{7}^{\boldsymbol{lookback}})$
Impractical to scale the algorithm beyond 7 *lookbacks*

How to improve memory efficiency for longer sequences?

Stock Movement as a Markov Process

3 ML-based predictor

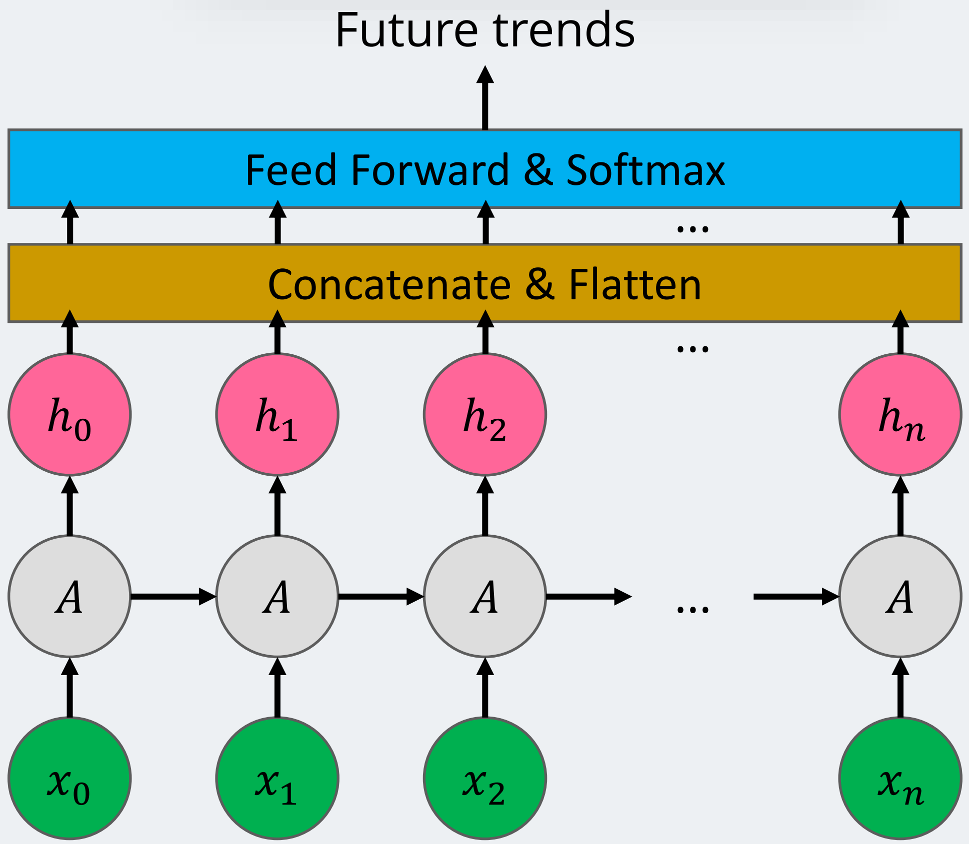
Recurrent Neural Network



Stock Movement as a Markov Process

3 ML-based predictor

Model Architecture



Stock Movement as a Markov Process

4 Use prediction in the strategy

$$P[X_{t+1} = (x_{t+1}, x_{t+2}, \dots, x_{t+forward}) \mid X_t = (x_{t-lookback+1}, x_{t-lookback+2}, \dots, x_t)]$$

The prediction of future trend is the price state with the highest probability

$$\hat{X}_{t+1} = \mathrm{argmax}_x \left(P[X_{t+1} = x \mid X_t = (x_{t-lookback+1}, x_{t-lookback+2}, \dots, x_t)]\right)$$

Experiment

1 Experimental setup

The historical price of stocks in S&P500 were collected
Total number of obtained stocks: 496

The stocks’ prices are categorized into 7 categories

Each data point is then one-hot encoded

Experiment

2 Model hyperparameters

```
self.model = SimpleLSTM(
    lookback=lookback,
    hidden_size=16,
)
```

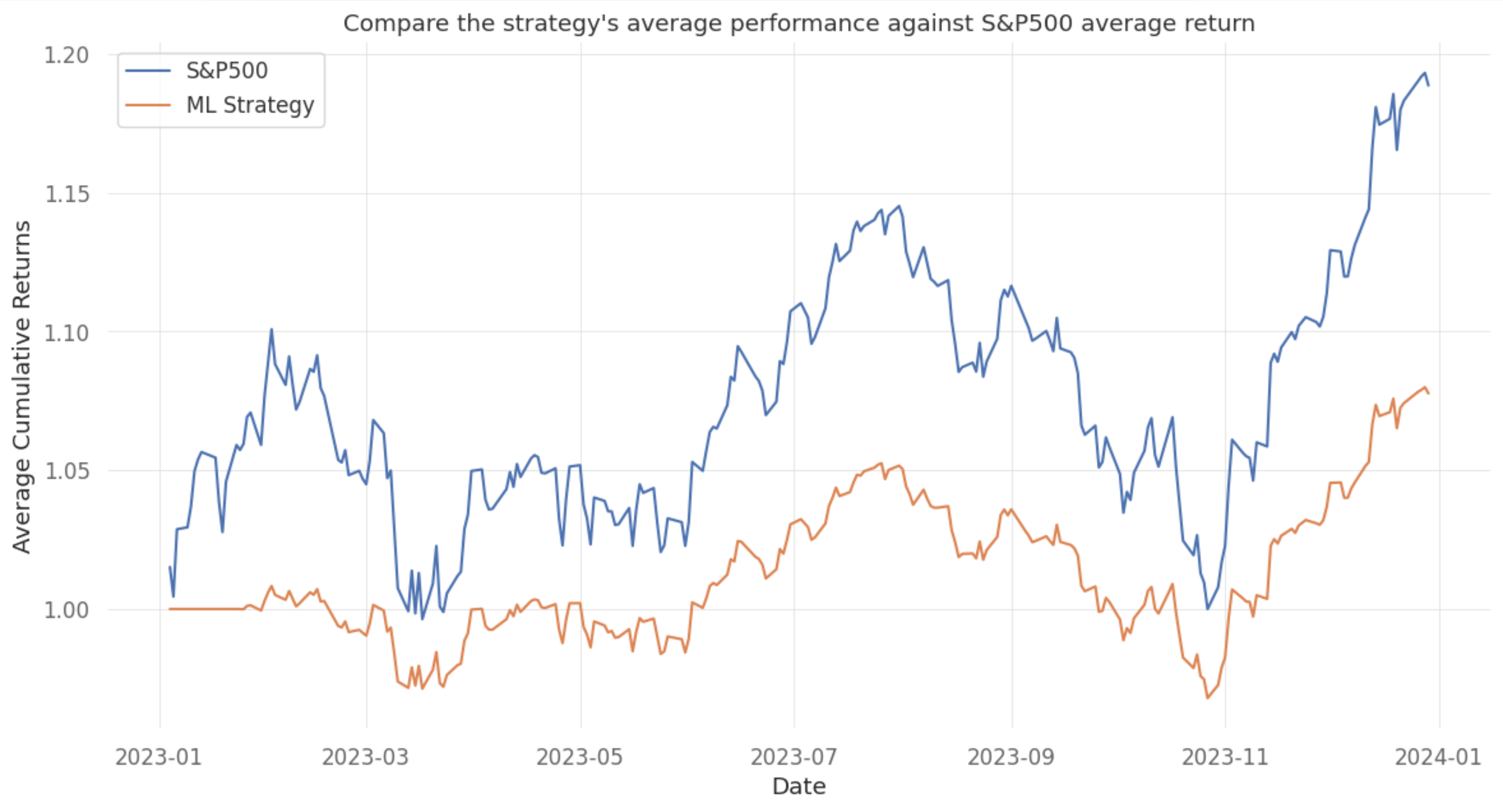
```
ml_strategy.train(
    epochs=100,
    lr=1e-2,
    batch_size=128,
)
```

$$Loss(y,\hat{y}) = -\sum_{x \in \{\text{up, side, down}\}} y_i \log(\hat{y}_i)$$

The model’s parameters are optimized by *AdamW optimizer*

Experiment

3 Experimental results



S&P500 Performance

Monthly Expected Return: 0.010
Monthly Volatility: 0.045
Sharpe ratio: 0.222

Strategy Performance

Monthly Expected Return: 0.006
Monthly Volatility: 0.025
Sharpe ratio: 0.226

Can it beat the market?

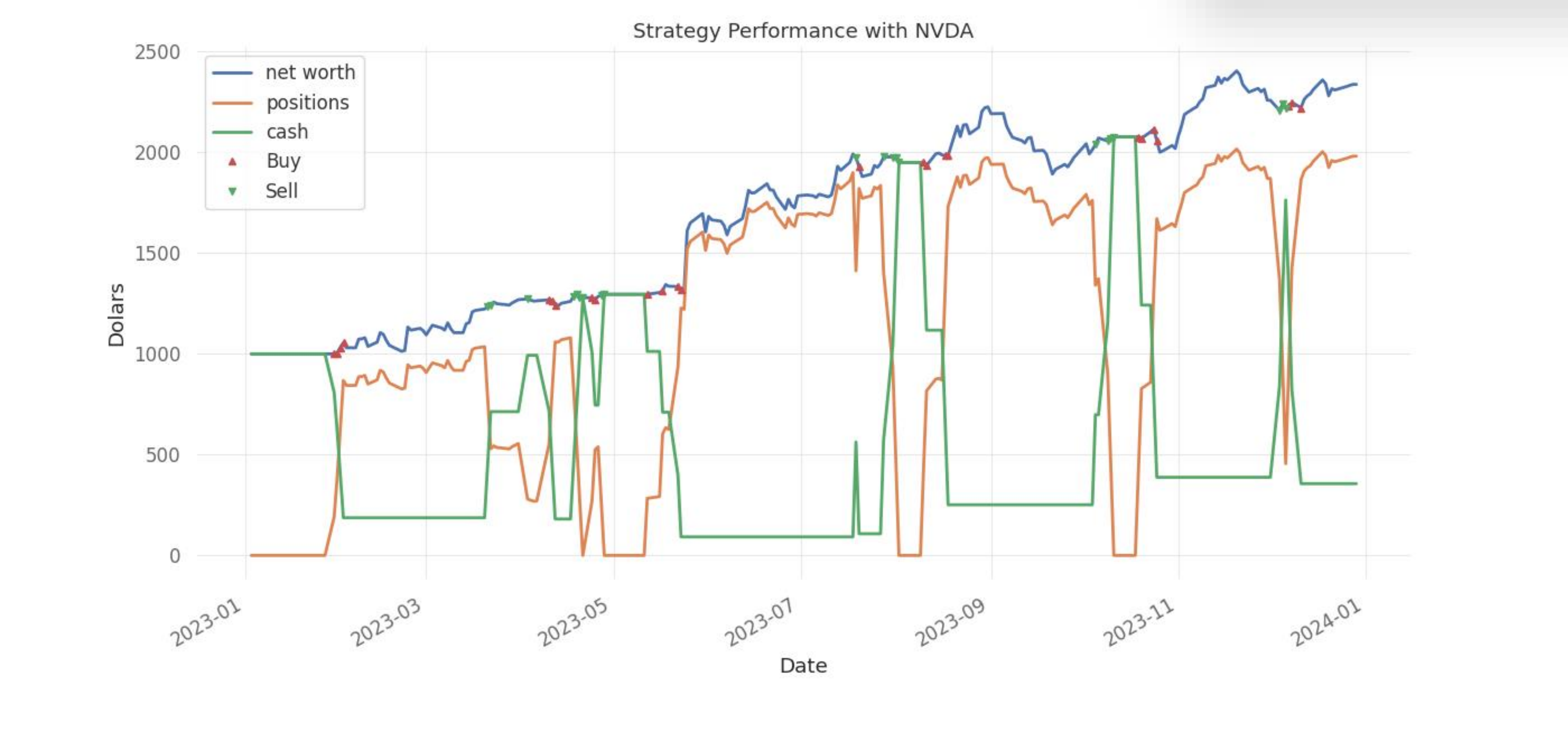
The strategy beats 243 stocks
→**Yes, but only 49% of the time**

The strategy generates higher returns than 196 stocks
→**Actually beat the market 49.52% of the time**

Experiment

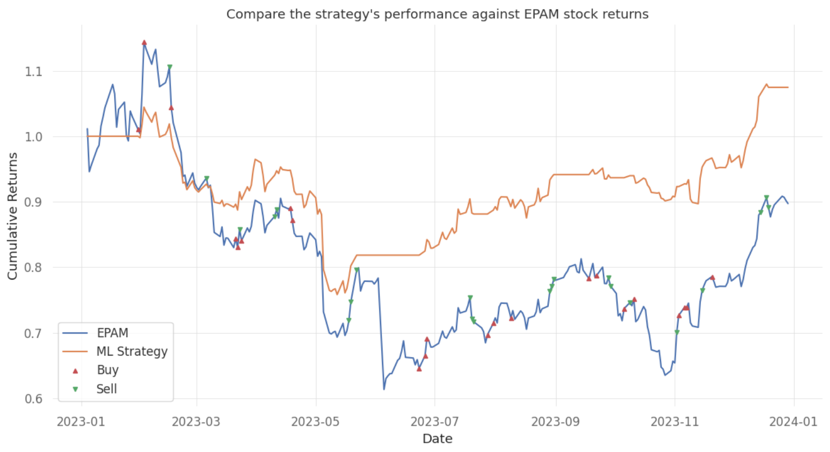
4 Case study - NVDA

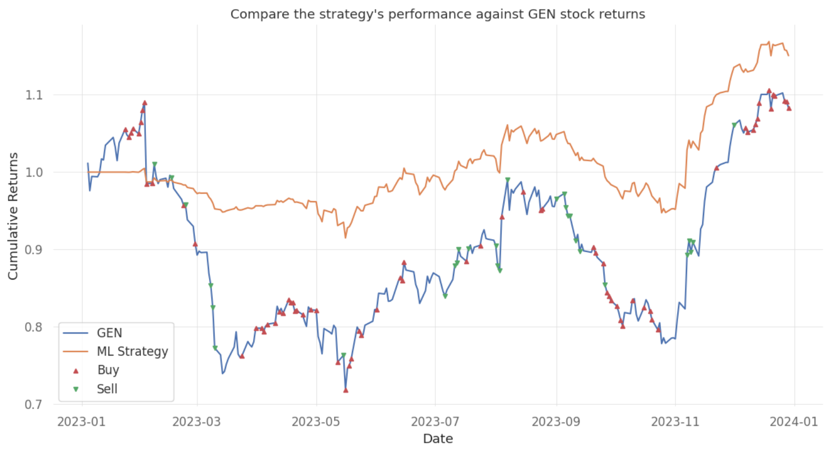
Monthly Expected Return: 0.082
Monthly Volatility: 0.094
Sharpe ratio: 0.874

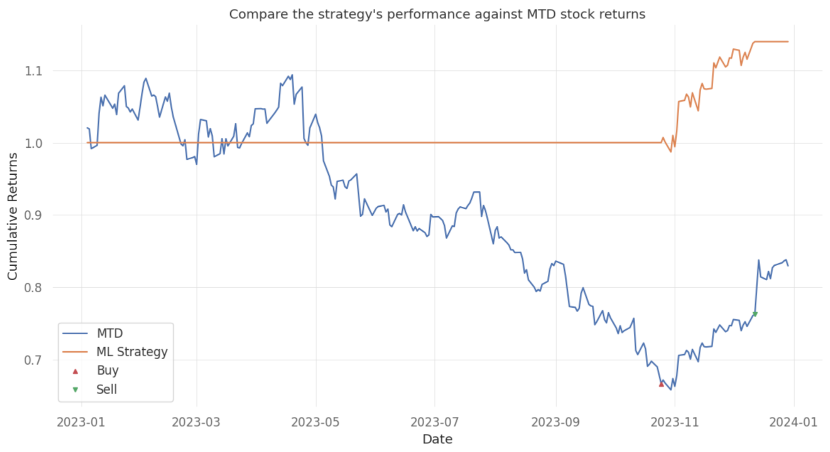


Experiment

4 Case study – Bearish Markets



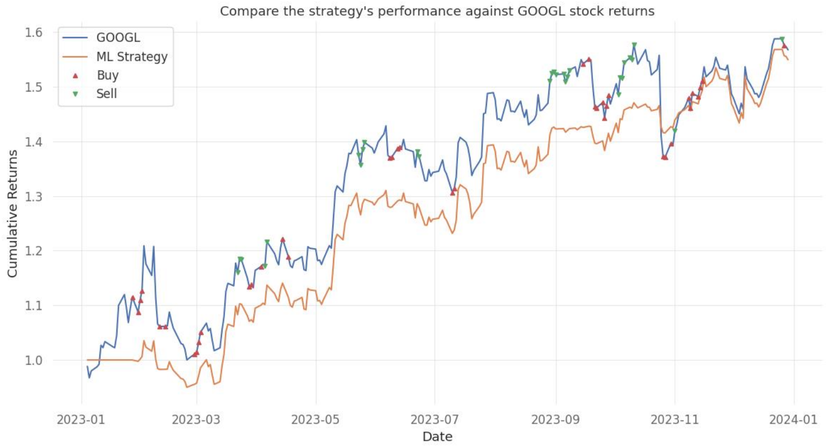


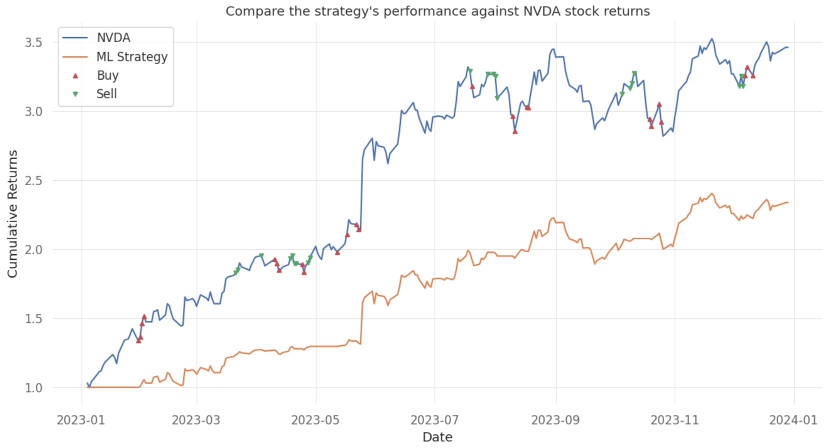


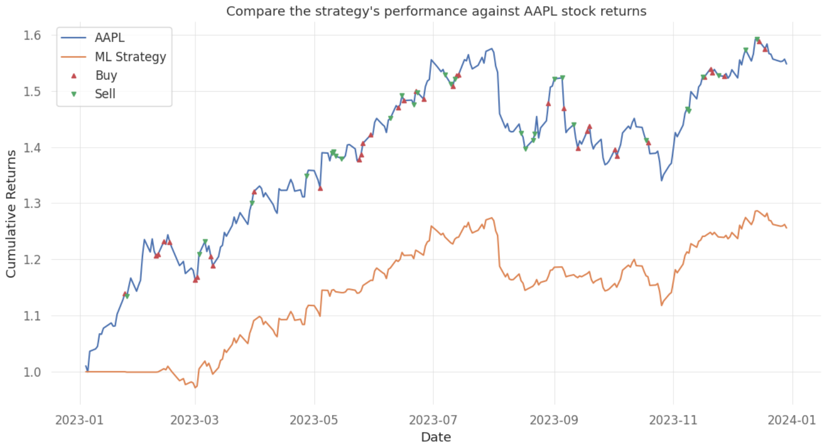
Experiment

4

Case study – Bullish Markets







Thank you