Algorithmic Trading Coursework 2

Student ID: 23075753

Department of Computer Science, University College London

March 20, 2024

1 Time Series Preparation

This report aims to design and evaluate trading strategies for the SPDR Portfolio Long Term Treasury ETF (SPTL), utilising data from Yahoo Finance spanning January 1, 2014, to December 31, 2019. To more accurately mirror actual trading conditions, the Effective Fed Funds Rate (EFFR Index) from the Federal Reserve Economic Data (FRED) for the same period is integrated, serving as the risk-free rate. Several data preprocessing steps are required to be conducted before analysis.

Deal with missing data When merging the dataset of SPTL and EFFR based on the timestamp, some rows might contain missing data. As trading occurs on weekdays, data from weekends are excluded. Additionally, on federal holidays, despite the absence of EFFR data and ongoing market activity, the EFFR value is forward-filled, based on the assumption that the rate on a holiday is identical to the previous trading day's rate.

Compute daily risk-free rate The conversion of the annual EFFR data, originally in percentage format, into a daily risk-free rate is crucial. This transformation involves dividing the annual rate by 25,200 (i.e., 252×100), where 252 refers to the typical number of trading days in a year, resulting in r^f, the daily risk-free rate.

Compute excess return per unit of SPTL The daily return of SPTL is the daily percentage change of the closed price, $r_t^d = \frac{p_t - p_{t-1}}{p_{t-1}}$. The excess return requires the daily EFFR rate to be subtracted from the daily rate: $r_t^e = r_t^d - r_t^f$. Figure 1 illustrates the time series plot for the SPTL daily return, EFFR and excess return.

2 Trading Strategies

In this section, 3 trading strategies are designed, starting with an initial capital of $V_{t=0} = \$200,000$ with a maximum leverage L of 10. At time t, the dollar value of the ETF

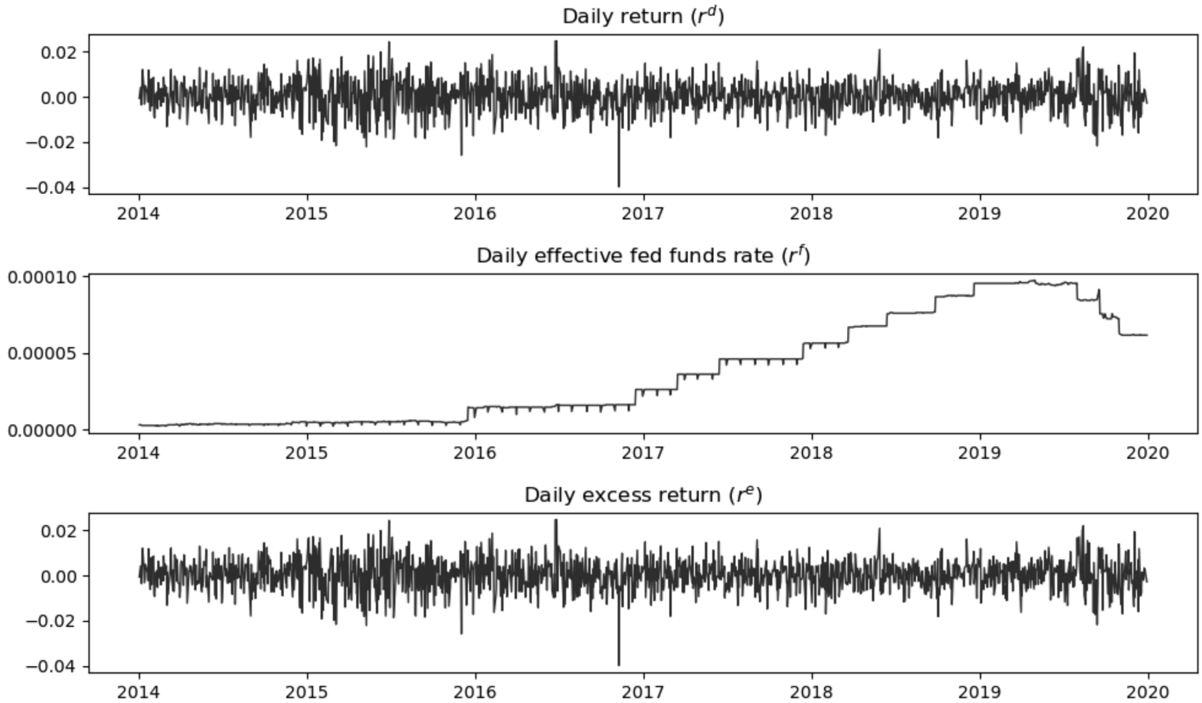

Figure 1: SPTL return time series, the EFFR and the excess return

invested (θ_t) must adhere to the constraint: $|\theta_t| \leq V_t \times L$. When the θ_t, surpasses the upper limit, the excess positions are liquidated at the current market price, with proceeds allocated to a money market account with interest at the risk-free rate. If θ_t, falls below the lower limit and no additional funds are available, trading activities must cease. The total portfolio value V^{total}, comprises V, the value of the holdings and V^{cap}, the unused capital invested at the risk-free rate (r_t^f). The trading strategies will first be designed based on the training set, which is the first 70% days of data. Their performance will be evaluated on the remaining 30% of the testing set data.

Before any trading activities, 95% of the initial $200,000 is designated for active trading (represented as V_0), while the remaining 5% is kept in V_0^{cap} to grow at the risk-free rate. This 5% reserve, amounting to $10,000, serves as a contingency fund. In scenarios where ETF short positions dip below the lower limit ($-V_t \times L$), this reserved fund in V^{cap} can be utilised to repurchase some ETF shares, adhering to the investment constraints.

2.1 Strategies Design

2.1.1 Moving Average and Relative Strength Index (RSI)

The paragraph presents a trading strategy integrating two Simple Moving Averages (SMA) and the Relative Strength Index (RSI) , emphasising its trend-following nature. The SMA, calculated over rolling time windows, establishes average price trends [1]. The RSI, comparing upward and downward price movements over 14 days, identifies over-bought (RSI above 70) and oversold (RSI below 30) conditions [2].

A buy signal is generated under two specific conditions: first, when the 30-day SMA crosses above the 50-day SMA, indicating that the short-term market trend is bullish; and second, if the RSI is below 50, which suggests that the asset is not yet overbought and has potential for an upward movement. This combination signals a favourable moment to enter a long position. Conversely, a sell signal arises from an opposite set of criteria: it occurs when the 30-day SMA moves below the 50-day SMA, indicating a bearish short-term trend. This signal is confirmed if the RSI is above 50. The choice of an RSI threshold at 50 is a practical adjustment that serves as a median level to identify market trends. The buy and sell signals generated can be shown in Figure 2.

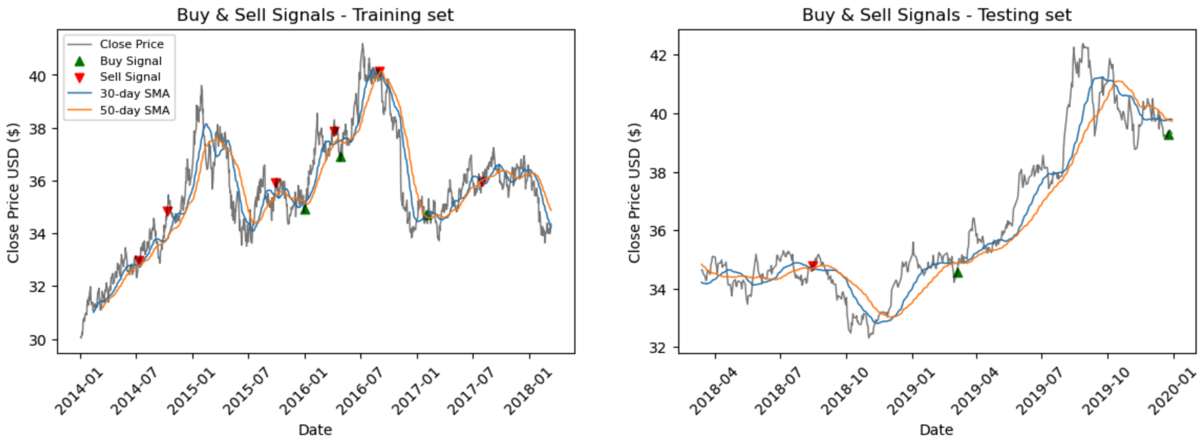

Figure 2: Buy and sell signals generated by SMA and RSI indicators

2.1.2 Moving Average Convergence Divergence (MACD)

The MACD is a trend-following strategy that can identify the direction and momentum of market trends. This indicator is constructed using two exponential moving averages (EMAs): the 12-day EMA and the 26-day EMA. Unlike SMAs, EMAs assign greater significance to recent data [3]. The MACD line represents the difference between these two EMAs. A zero value in the MACD line indicates a crossover of the 12-day and 26-day EMAs. Furthermore, the MACD strategy incorporates a signal line, which is the 9-day EMA of the MACD line itself [4]. A bullish signal is suggested when the MACD line crosses above the signal line, indicating a long signal. Conversely, a bearish signal is inferred when the MACD line crosses below the signal line, suggesting a shorting opportunity. The buy and sell signals generated by this strategy are shown in Figure 3.

2.1.3 Long Short-Term Memory (LSTM) Model

Long Short-Term Memory (LSTM) models, a type of Recurrent Neural Network, are well-known for their ability to predict sequential data like time series, making them well-suited for financial market analysis [5]. LSTMs excel in capturing both short-term and long-term trends in data sequences. The core of LSTM's design lies in its cell state and gating mechanisms. These include the input, output and forget gates, which collectively

Figure 3: Buy and sell signals generated by the MACD indicator

manage the flow of information. They allow the model to retain or discard information, making it adept at learning from historical trends to forecast future events.

The LSTM model used in this strategy is built using Keras and Tensorflow in Python to predict future closed prices. Given the price history of the previous 6 days, the model is trained to predict tomorrow's closed price. Its architecture includes two LSTM layers of 100 and 50 neurons, a 25-neuron dense layer, and a final output layer. The model employs the Adam optimizer to minimize mean squared error (MSE). The training involved 30 epochs with a batch size of 15. The MSE on the training set is 2.609 and it increases to 3.571 on the testing set, suggesting less accurate predictions for unseen data, as depicted in Figure 4. This discrepancy between training and testing performance highlights potential overfitting issues. Due to the non-deterministic nature of deep learning, the price prediction for each execution might differ. The selected architecture has undergone extensive testing and validation to guarantee reliability in its price prediction capabilities.

A buy signal is generated if the predicted price for the next day exceeds today's price by 2%. Conversely, a sell signal is created if the predicted price is 2% lower than the current day's price.

2.2 Position and Turnover

The three strategies are executed using a straightforward yet aggressive approach. Constantly employing a constant maximum leverage of 10, the portfolio entirely shifts to a long position with every buy signal, and conversely, moves to a full short position upon each sell signal. When no trading signals are received, the previous positions will hold. The changes of positions in SPTL of the three strategies are shown in Figures 5 to 7,

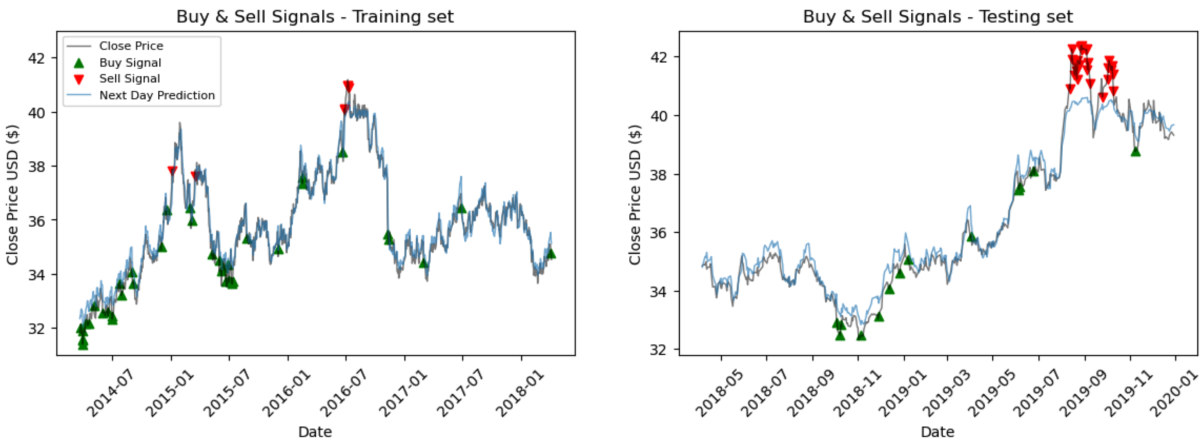

Figure 4: Buy and sell signals generated by the LSTM strategy

which are bounded by the asset value (V_t) times the leverage of 10.

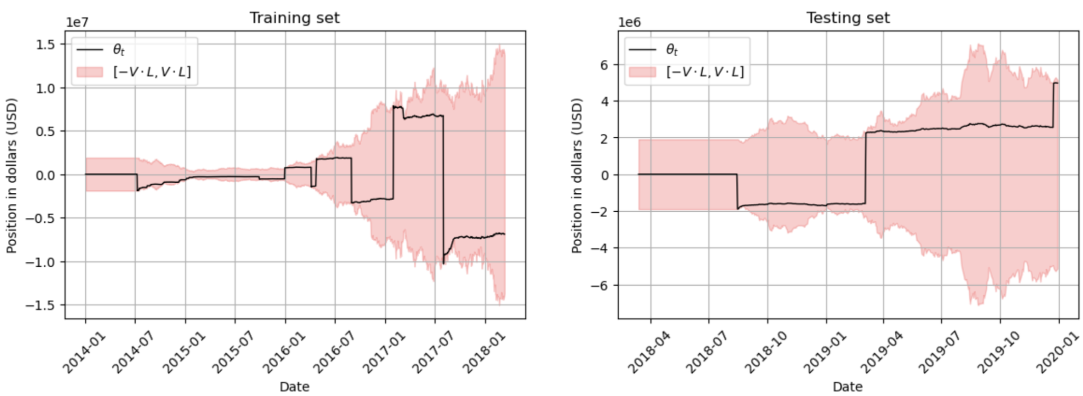

Figure 5: Position plot of the MA & RSI strategy

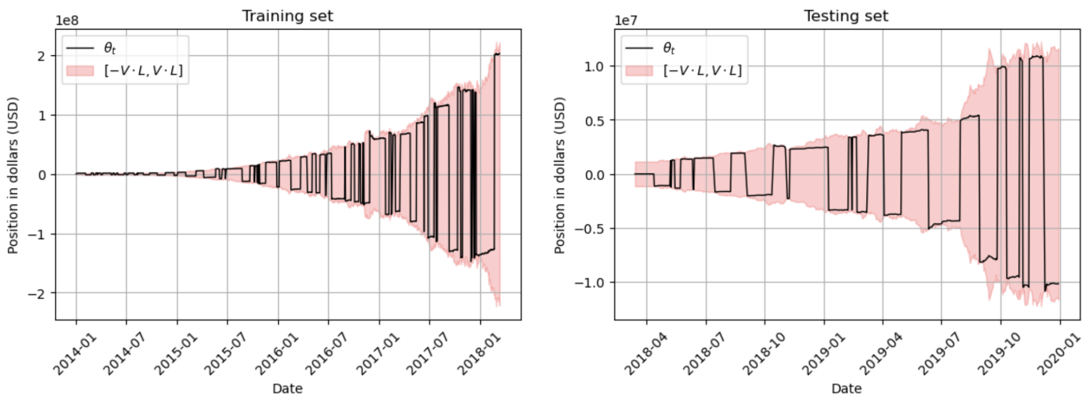

Figure 6: Position plot of the MACD strategy

Turnover is a measure of trading activity, that can be represented in dollar value or units traded, according to the equations below, where T is the time window and the p_t is the closed price of the ETF at time t.

$$Turnover_{dollars} = \sum_{0}^{T} |\Delta\theta_t|$$

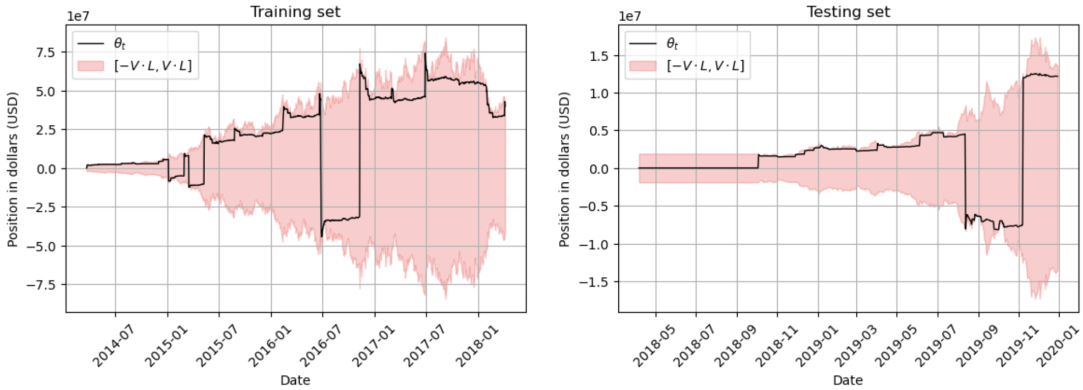

Figure 7: Position plot of the LSTM strategy

$$Turnover_{units} = \sum_{0}^{T} |\frac{\theta_{t+1}}{p_{t+1}} - \frac{\theta_t}{p_t}|$$

A high turnover in a trading strategy indicates frequent trading activities. This often results in increased transaction costs in practical trading scenarios. Figures 8 and 9 demonstrate the 30-day moving average of turnover, measured in dollars, for each of the three strategies during both the training and testing phases. The turnover calculated in units is not shown, as its trends closely resemble those of the dollar-based turnover, due to its derivation: $\theta_t = units(t) \times p_t$. The total sums of the turnover in both the training and testing set are presented in Table 1.

The MA & RSI strategy exhibits the lowest frequency of trades, while the MACD strategy is the most active one. It's observed that turnover increases during periods of high price volatility, as significant price shifts often trigger the strategy's indicators to identify trading opportunities. Notably, during periods like the third and fourth quarters of both 2016 and 2019, when price fluctuations were prominent, a corresponding rise in turnover spikes is evident in the charts.

Strategy	MA & RSI		MACD		LSTM	
Dataset	Train	Test	Train	Test	Train	Test
Turnover (Dollars)	60,856,895	12,230,681	7,456,339,726	262,432,381	626,231,524	59,250,576
Turnover (Units)	1,342,437	234,323	200,905,162	6,728,297	12,978,130	1,356,764

Table 1: Summary of trading strategy turnover

2.3 Profit and Loss

The trading strategies devised from the training dataset are subsequently assessed on the testing set to evaluate their performance with unseen data. The daily profit and loss (PnL) of the asset value is ΔV_t, which is the difference of V_t and V_{t-1}. Similarly, the daily PnL of the unused capital is ΔV_t^{cap}. The total daily profit is calculated based on the equation below, where the ratio $\frac{|\theta_t|}{L}$ is known as the margin, M_t. By summing up all the

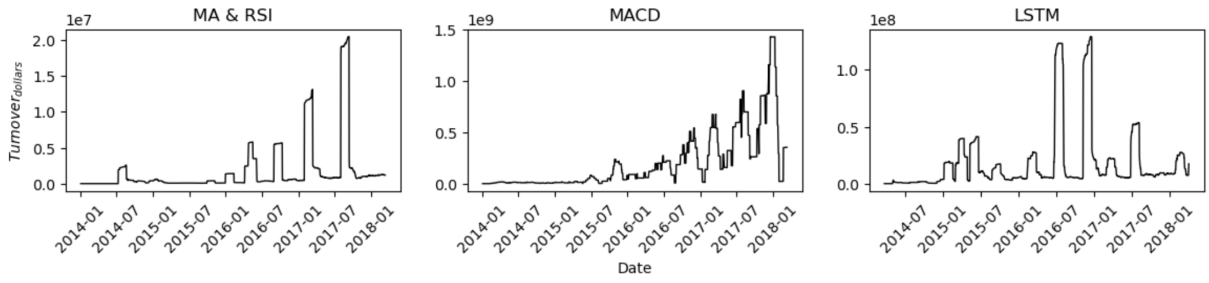

Figure 8: Turnover of the training set (30-day window)

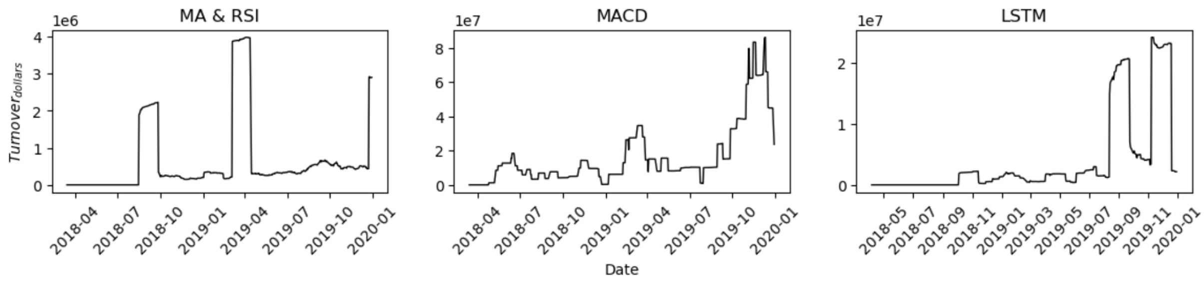

Figure 9: Turnover of the testing set (30-day window)

total daily profit, the cumulative profit of the strategy can be computed.

$$V_{t+1}^{total} - V_t^{total} = \Delta V_t^{total} = \Delta V_t + \Delta V_t^{cap} = (\frac{\Delta p_t}{p_t} - r_t^f)\theta_t + (V_t^{total} - \frac{|\theta_t|}{L})r_t^f$$

As shown in Figures 10 and 11, the MACD strategy is associated with both high risks and high rewards. Although it exhibits the most significant daily PnL swings, it ultimately achieves the highest cumulative profit at \$1,746,361 on the testing set, as detailed in table 2. Conversely, the MA & RSI strategy, which conducted the fewest trades, yields the smallest profit at \$325,224 on the testing set. It's important to note that these PnL calculations do not account for transaction costs and the potential costs of the high-leverage usage. Actual profits for the MACD strategy are expected to be lower when adjusting for the higher turnover and associated transaction expenses.

Strategy	Training Set Cumulative PnL (Dollars)	Testing Set Cumulative PnL (Dollars)
MA & RSI	1,167,103	325,224
MACD	36,734,603	1,746,361
LSTM	4,125,483	1,270,093

Table 2: Profit and loss on training and testing set

The risk-free rate has a direct impact on the total profit based on the equation above. If r_t^f increases, the daily excess return $(\frac{\Delta p_t}{p_t} - r_t^f)$ will decrease, resulting in a lower ΔV_t. However, with a higher r_t^f, the earning on V_t^{cap} will increase. The influence of r_t^f on the total PnL depends on the portfolio allocation of each strategy. By setting a 150% r_t^f, the cumulative PnL of the MA & RSI strategy will increase by 1.51%, while the MACD and LSTM strategies' PnLs will decrease -0.29% and -6.85% respectively. These results

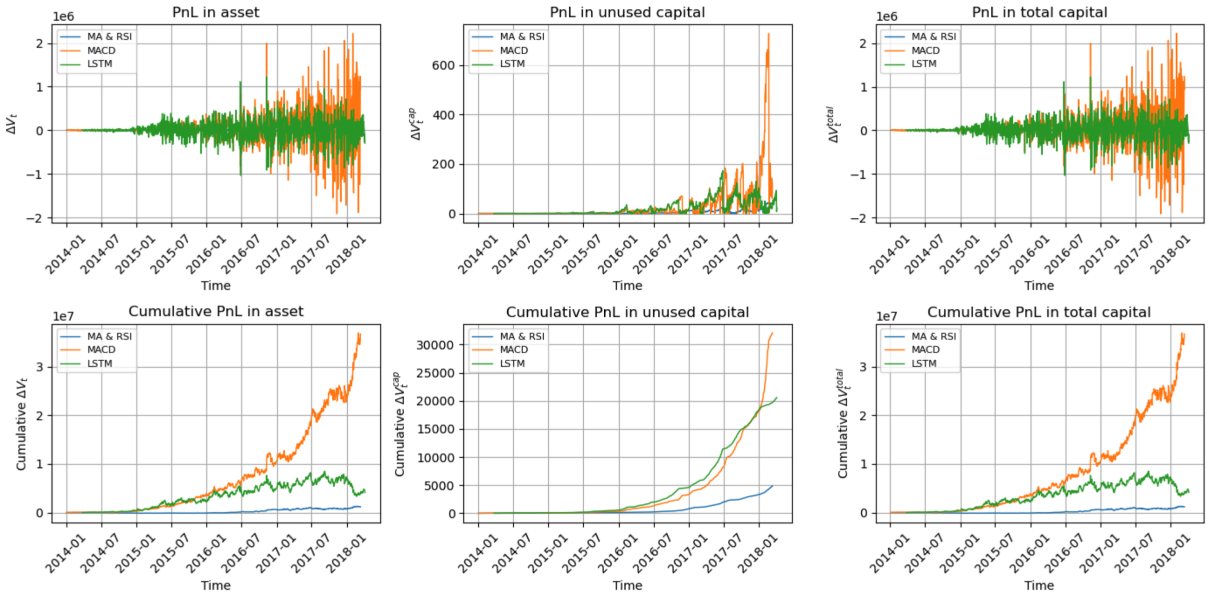

Figure 10: Profit and loss on the training set

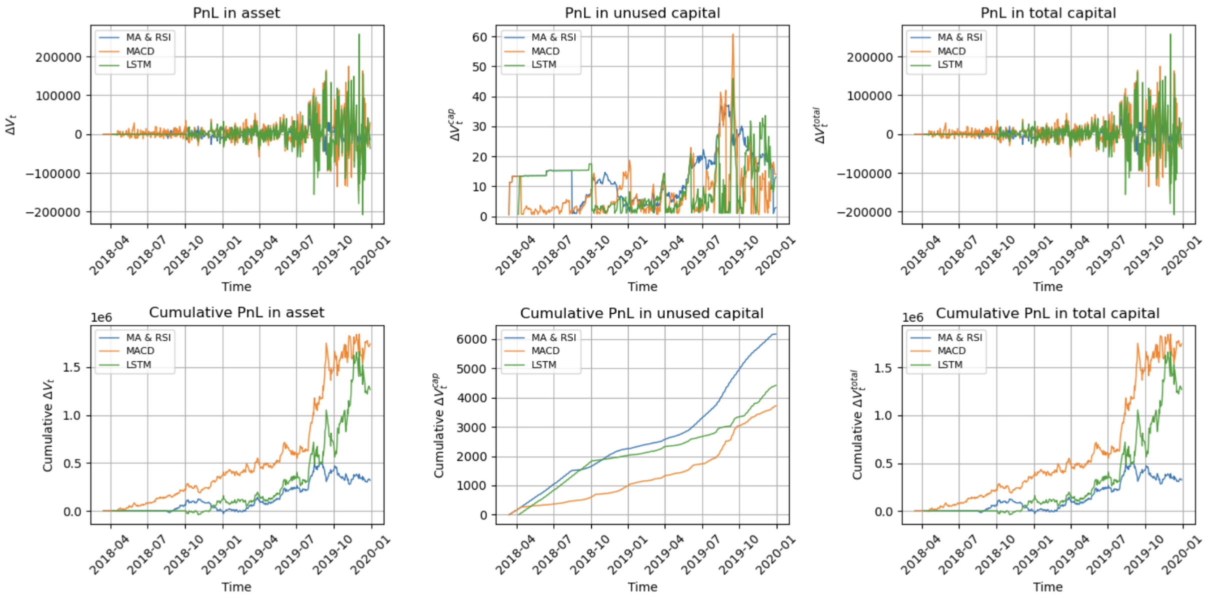

Figure 11: Profit and loss on the testing set

suggest that for the MACD and LSTM strategies, the reduction in excess return on ΔV_t outweighs the earnings on ΔV_t^{cap}, while MA & RSI strategy is the opposite.

3 Performance Indicators

3.1 Common Ratios

Sharpe Ratio (SR): The Sharpe Ratio is an assessment of risk-adjusted return, comparing the excess return of an investment to its volatility [6]. The ratio is defined with the formula below, where r_p refers to the portfolio's return, r_f is the risk-free rate and σ_p

is the standard deviation of the excess return. To annualise the ratio, a factor of $\sqrt{252}$ will be multiplied, as there are 252 trading days a year.

$$\text{Sharpe Ratio} = \frac{r_p - r_f}{\sigma_p}$$

Sortino Ratio: The Sortino Ratio modifies the Sharpe Ratio by focusing on downside risk. It uses σ_d instead, the standard deviation of negative asset returns, instead of the total standard deviation [7]. This ratio penalises only the negative volatility, assuming investors are primarily concerned with the risk of losses. The same rule will be applied to get the annualised ratio, as the Sharpe Ratio.

$$\text{Sortino Ratio} = \frac{r_p - r_f}{\sigma_d}$$

Maximum Drawdown (MDD): Maximum Drawdown represents the largest peak-to-trough decline in an investment's value before a new peak is achieved [6]. It is often reported as a percentage and provides investors with an indication of potential losses during a specified investment period.

$$\text{Maximum Drawdown} = \frac{\text{Min Trough Value} - \text{Max Peak Value}}{\text{Max Peak Value}}$$

Calmar Ratio (CR): The Calmar Ratio evaluates the performance of an investment relative to its downside risk. It is calculated by dividing the annual rate of return (μ) by the maximum drawdown defined above. A higher CR suggests that the investment has yielded more return per unit of risk, representing a more efficient performance.

$$\text{Calmer Ratio} = \frac{\mu}{MDD}$$

Strategy	MA & RSI		MACD		LSTM	
Dataset	Train	Test	Train	Test	Train	Test
Sharpe Ratio	0.50078	0.95890	3.60467	2.84050	0.69003	1.49914
Sortino Ratio	0.74062	1.32876	5.82655	5.03477	0.98948	1.92036
Maximum Drawdown	-0.90563	-0.52381	-0.28939	-0.28420	-0.88804	-0.43324
Calmar Ratio	0.85146	1.35565	15.48069	8.95883	1.48036	3.50385

Table 3: Summary of performance indicators

Table 3 presents an evaluation of three distinct trading strategies using the defined performance indicators. Both MA & RSI and LSTM strategies exhibit improvement in the test set over the training set, as indicated by the increased ratios, a likely consequence of favourable market trends. These strategies, taking fewer positions, benefit from prolonged holds during an uptrend, compared to the MACD strategy that buys and sells more frequently. The MACD strategy has significantly higher Calmar and Sortino ratios, indicating strong performance during the training phase. However, a performance decline is seen in the testing set, due to potential overfitting issues.

3.2 Rolling Sharpe Ratio

The 90-day rolling SR in Figure 12 illustrates changes in strategies' performances over time. The rolling SR of MACD is maintained more consistently at a positive level with fewer drawdowns, compared to other strategies, despite market up or down trends. In contrast, the MA & RSI and LSTM strategies' SR are more influenced by market conditions, reflecting their longer holding periods and less frequent trading. In comparison, the MA & RSI is less robust as it has a larger drawdown of SR observed in both training and testing sets. To counteract the disparity between training and test set performances, methods like out-of-sample testing are advisable. Such methods validate strategy robustness and ensure that the model captures patterns that are generalisable to new, unseen data and prevent the strategy from being too dependent on a particular market trend.

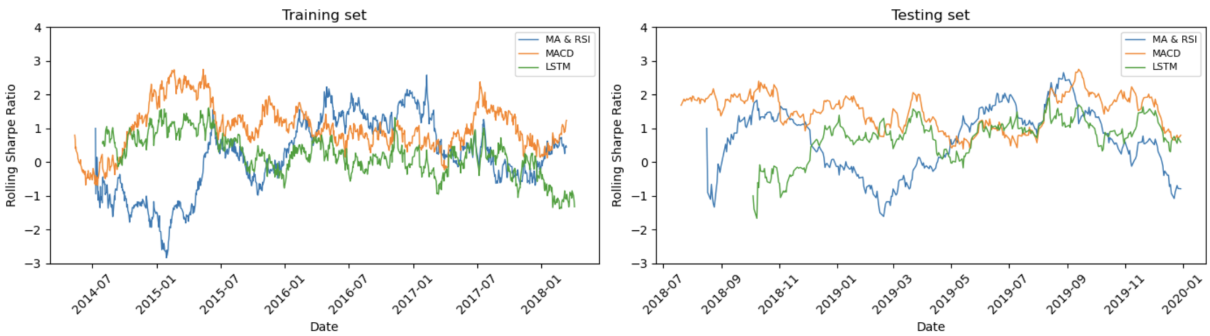

Figure 12: 90-day rolling Sharpe ratio of three strategies

3.3 Profit and Loss Drawdown

Figures 13 to 15 show the drawdowns of the cumulative profits with the market volatility, which is calculated as the 90-day rolling standard deviation of the ETF returns. Throughout the training period, a marked increase in drawdown is observable across all three strategies. This trend is mainly caused by the accumulation of profits, a consequence of the strategies' high-leverage, aggressive trading approach. The trading activities notably amplify profits towards the end of the period, thereby elevating the risk of substantial drawdowns, independent of ETF volatilities. The MA & RSI approach exhibits a prolonged drawdown phase spanning from July 2014 to January 2016, coinciding with a volatility peak, as depicted in Figure 13. A similar but smaller drawdown pattern could be observed in the LSTM strategy as well. In the test dataset, all strategies experience significant drawdown towards the end, particularly as volatility peaks. The MA & RSI strategy exhibits greater drawdown patterns during the periods when volatility spikes, especially in January and April 2019. Overall, the volatility seems to have the least impact on the MACD strategy, suggesting a potentially robust approach less influenced by volatility fluctuations.

Generally, in periods of high market volatility, prices can fluctuate widely in a short amount of time. Using less margin during these times is considered a safer strategy because it reduces the potential for large losses, resulting from sudden market movements against the current position. Conversely, when the volatility is low, when price movements are smaller and more predictable, it might be safer to use more margin. A great strategy is to adjust the leverage dynamically based on the underlying price volatility. In addition, incorporating stop loss and take profit orders helps to better manage the risks associated with leveraged positions.

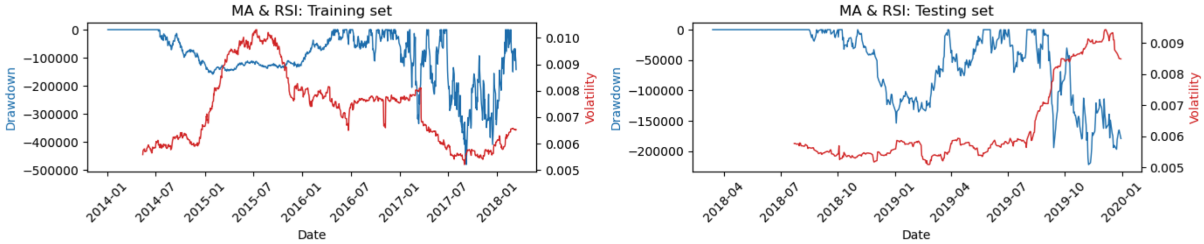

Figure 13: Drawdown and volatility plot of MA & RSI strategy

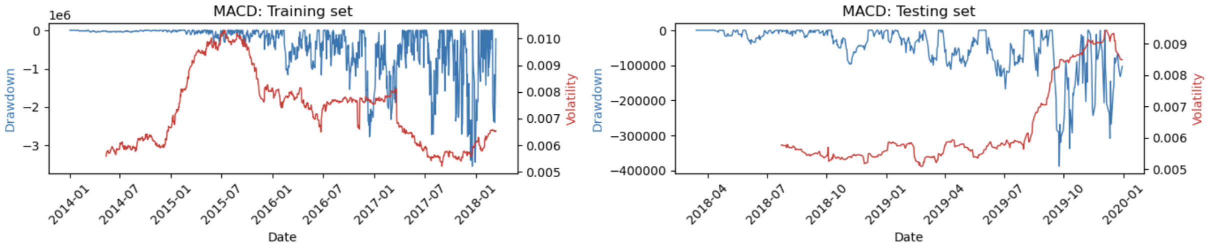

Figure 14: Drawdown and volatility plot of MACD strategy

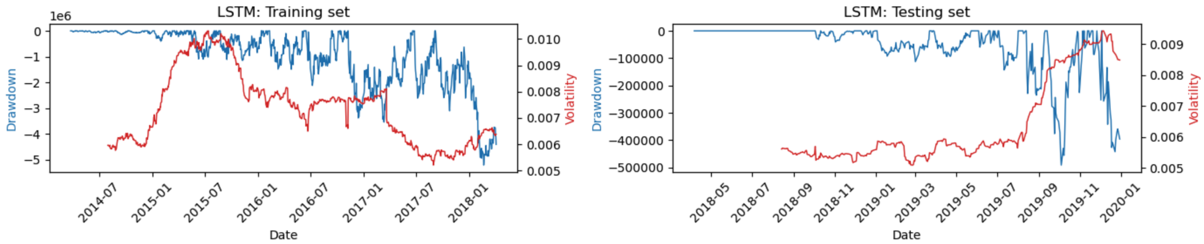

Figure 15: Drawdown and volatility plot of LSTM strategy

4 Discussion and Conclusion

This report has developed three distinct leveraged trading strategies utilising data from the SPTL ETF. In the strategy formulation phase, these three strategies are identified as superior in performance relative to others not discussed in this report, such as mean reversion, Bollinger Bands, auto-regressive models, and a simple buy-and-hold approach. Among the strategies analysed, the MACD strategy is the top performer, yielding the highest profits across both the training and testing datasets. It also has the highest

Sharpe, Sortino and Calmar ratios, indicating a superior risk-adjusted return. Moreover, the MACD strategy demonstrates robustness through its consistent performance, showing little correlation with market volatility. In contrast, the MA & RSI strategy is the least effective, with its performance heavily influenced by market trends despite its straightforward implementation.

Several enhancements are suggested to improve these strategies. One area for improvement is the dynamic adjustment of leverage and position sizing upon receiving buy or sell signals. Currently, all three strategies employ a maximum predefined leverage level and commit the entire capital to each trade, irrespective of whether it involves taking a long or short position. This method is aggressive and carries significant risk. Implementing portfolio optimisation techniques to fine-tune position sizes could mitigate this risk. Additionally, while some trading entities may avoid the risk associated with prolonged position holding, the strategies described do not limit the duration for which a position can be held. Incorporating mechanisms to preferably shorten the holding period could further align these strategies with the risk preferences of more conservative trading firms.

References

[1] Adam Hayes. *Simple moving average (SMA): What it is and the formula*. Sept. 2023. URL: https://www.investopedia.com/terms/s/sma.asp.

[2] Jason Fernando. *Relative strength index (RSI) indicator explained with formula*. Feb. 2024. URL: https://www.investopedia.com/terms/r/rsi.asp.

[3] James Chen. *What is Ema? how to use exponential moving average with formula*. Feb. 2024. URL: https://www.investopedia.com/terms/e/ema.asp.

[4] Brian Dolan. *What is MACD?* Mar. 2024. URL: https://www.investopedia.com/terms/m/macd.asp.

[5] Shipra Saxena. *What is LSTM? Introduction to long short-term memory*. Jan. 2024. URL: https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/.

[6] John. *Sharpe, Sortino and Calmar Ratios with Python*. Oct. 2020. URL: https://www.codearmo.com/blog/sharpe-sortino-and-calmar-ratios-python.

[7] Will Kenton. *Sortino ratio: Definition, formula, calculation, and example*. Mar. 2024. URL: https://www.investopedia.com/terms/s/sortinoratio.asp.