.

Aim: To optimize trading strategies using DRL, NOT price prediction or portfolio management

Research Objectives:

I. To improve the performance of stock trading models by applying feature engineering techniques to obtain technical indicators used by DRL models.
II. To improve the performance of stock trading models by using DRL during normal stock market conditions in simulated trading environments with performance metrics such as Cumulative Return and Sharpe Ratio.
III. To improve the performance of stock trading models by using DRL during bearish and volatile stock market conditions in simulated trading environment with Maximum Drawdown as a performance metric.

Datasets Used: AMZN, MSFT, GOOGL historical stock datasets of 10 years (2013-2023).

Parts Already Done: Data Cleaning, Preprocessing, Feature Engineering, EDA, Feature Selection and Recursive Feature Elimination.

Model Construction:

1. Trading Environment
2.

Table 4.4. Components of the Trading Environment.

Components	**Description**
State Space (S)	Open, High, Low, Close, SMA, EMA, MACD
Action Space (A)	Sell (0), Hold (1), Buy (2)
Reward Function (R)	$Selling\ Price - Buying\ Price$
Transfer Probability (T)	Implicitly learned by the DQN model.
Discount Factor (γ)	0.95

3. DRL Model: **DQN with LSTM** neural network architecture.
4. Model Training.

Results Analysis:

1. Cumulative Return

$$Cumulative\ Return\ , CR = \frac{Final\ Value - Initial\ Value}{Initial\ Value} \times 100 \qquad (3,8)$$

where the initial value is the initial value of the portfolio, and final value is the final value of the portfolio (Yang H et al., 2020).

2. Sharpe Ratio

$$Sharpe\ Ratio, SR = \frac{Rp - Rf}{\sigma p} \qquad (3,9)$$

where Rp is the return, Rf is the risk-free rate, and σp is the standard deviation of excess returns (Yuan Y et al., 2020).

3. Maximum Drawdown

$$Maximum\ Drawdown, MDD = \frac{Peak\ Value - Valley\ Value}{Peak\ Value} \times 100 \qquad (3,10)$$

where the peak value is the highest value before the largest drop and valley value is the value before a new high value is established (Carta S et al., 2021; Yu X et al., 2023).